Chapter 1

INTRODUCTION AND OVERVIEW

Economic policies are typically based on the assumption that actors do not cooperate voluntarily if it involves personal costs. This assumption is central in neoclassical economic theory where actors' behavior is attributed to the exclusive maximization of self-interest. Economic theory has provided accurate predictions for field behavior whenever individual interests are in conflict with collective interests and therefore *seems* to be an accurate guide for economic policies. Take the over-exploitation of common pool resources like fishing grounds as an example. Economic theory suggests that common pool resources are overexploited in the absence of sanctions; the actual depletion of fishing grounds in many places appears to corroborate the assumption that actors are self-interested and do not cooperate voluntarily to sustain common pool resources.

However, deducing that actors are self-interested from field data can be misleading. The problem is that the falsification of the standard assumption with field data is exceedingly difficult. On the one hand, observing frequent overexploitation of common pool resources can – but does not necessarily – imply that resource users have selfish preferences. On the other hand, observing some sustained common pool resources does not automatically imply that resource users do not have selfish preferences. Common pool resource exploitation, like most other field behavior, is explainable by all sorts of factors that are often uncontrollable and imperfectly observable. For instance, external factors like industrial pollution, an increase in the number of resource users, or climate change can affect resource exploitation, whereas the sustainable use of common pool resources can be due to social pressure, informal rules, or budget constraints that prevent the purchase of machinery that facilitates resource exploitation.

Economic policies can only be efficient if they are based on accurate assumptions. With respect to the management of common pool resources, the prescriptions derived from the standard assumption are typically to monitor the resource exploitation by external agents or to privatize the resources. These prescriptions are often either exceedingly expensive or infeasible which could lead to the impression that common pool resources are unmanageable in some cases and that their overexploitation is unavoidable. If, however, the standard model was inaccurate and at least some resource users were not selfish but cooperative, different ways for successful management of common pool resources could be feasible. For example, one alternative way is to empower resource users to self-manage and monitor the use of resources; this may work, given there are cooperative resource users. Therefore, increased knowledge about individuals' preferences and how they influence field behavior is of utter importance for designing accurate policy prescriptions.

One approach for rigorously testing individual preferences is with laboratory experimentation. Numerous laboratory experiments have shown extensive and consistent departures from the standard assumption (for reviews see e.g. Camerer, 2003 and DellaVigna, forthcoming). However, it is unclear how far laboratory behavior can be used as a guide for field behavior or even for the recommendation of economic policies. First, it cannot be taken for granted that laboratory behavior can be generalized to field behavior because the laboratory environment usually differs systematically in various domains to the environment outside the laboratory (Levitt and List, 2007). For instance, subjects in the laboratory know that an experimenter scrutinizes them, which could make them behave differently than in naturally occurring situations (e.g. Orne, 1962). Moreover, by abstracting from the situational complexities in the field, it is possible that important factors which determine field behavior to a large extent are neglected or cannot be simulated in the laboratory. Thus, simply extrapolating laboratory behavior to field behavior and suggesting policy recommendations based on laboratory findings seems premature.

One possibility for testing the relevance of laboratory behavior as a predictor for field behavior is by combining laboratory preference measures and field behavior from the same individuals. This approach can circumvent the problems mentioned above that occur if the researcher has only access to either laboratory or field data. In particular, if there are relationships between laboratory and field behavior, it provides evidence that laboratory behavior is to some degree generalizable to field behavior and it also helps understand the extent to which different individual preferences influence field behavior. The first two essays in this thesis apply this approach and provide evidence that there are indeed significant relationships between laboratory preference measures and field behavior and that the standard assumption is too narrow for understanding field behavior. In particular, they show that *cooperativeness* (which captures other-regarding preferences) and *impatience* (time preferences) play an important role in field behavior. The third essay in this thesis uses laboratory experimentation and a classification procedure to test competing explanations for *punishment* behavior in the laboratory.

It has been observed in numerous laboratory experiments that some actors are cooperative in the presence of a conflict between private and collective interests. In public goods games, for instance, we observe that many subjects spend money to contribute to public goods (Andreoni, 1995; Ledyard, 1995; Fehr and Gächter, 2000). It has also been observed that many actors are conditionally cooperative, i.e. they are cooperative, but only if they believe that other actors will cooperate as well (Fischbacher et al. 2001; Frey and Meier, 2004). These behavioral patterns in the laboratory appear to be robust to different subject pools and stake size (Cameron, 1999; Fehr et al., 2002; Fehr and Leibbrandt, 2008), and the findings give reason to believe that some actors have other-regarding preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Falk and Fischbacher, 2006).

There is also laboratory evidence that individuals have different time preferences and that many individuals are highly impatient (Benzion et al., 1989; Ashraf et al. 2006). Time preferences can have an important impact on behavior; for instance, it is likely that a wide range of empirically observable field behaviors, like low saving rates or high credit card take-up (e.g. Laibson et al., forthcoming), are due to impatience. There are already empirical studies which combine laboratory time preference measures with field behavior, showing that time preferences are actually directly related to financial behavior (Meier and Sprenger, 2007). Moreover, time preferences seem important for cooperation. The more impatient individuals are, the less they invest in cooperation for future outcomes like the conservation of scarce resources. To date there are, however, no empirical studies showing the direct link between time preferences and cooperation in general or resource exploitation in particular. The first two essays in this dissertation close this gap and provide evidence how time preferences are related to cooperation in the field.

A different frequent laboratory finding is that many subjects incur costs to punish other subjects even if no future benefits can be expected from this behavior. This propensity to punish altruistically posses a challenge to the standard assumption of self-interested individuals and can be of interest when designing economic policies. In public goods games with a punishment stage, for instance, it can be observed that many subjects punish free-riders and that this punishment behavior helps sustain high contribution levels to public goods (Ostrom et al., 1992; Fehr and Gächter, 2000; Masclet et al., 2003). Extrapolated to the field, this could imply that some individuals are willing to punish uncooperative behavior and in this way enforce cooperation norms. However, the investigation of laboratory punishment behavior is still in its infancy and the motivations behind punishment are not perfectly understood. In particular, it is unclear whether individuals punish in an impartial and normative manner, or whether they do so egocentrically. Thus, before giving any recommendations for economic policies (for instance the facilitation of informal punishment among resource users), increased knowledge about the motivations behind punishment is crucial, as is an examination whether punishment promotes desirable outcomes and deters undesirable outcomes or if it instead hampers desirable outcomes.

The goal of this thesis is to improve our knowledge about the relevance of cooperativeness, impatience, and punishment for economic outcomes. The first essay, "*Cooperativeness and Impatience in the Tragedy of the Commons*", examines the role of cooperativeness and impatience in the exploitation of common pool resources. In this study, we investigate fishermen whose main source of income stems from fishing grounds with open access. We combine laboratory measures of other-regarding and time preferences from the same fishermen with data about their fishing instruments. The findings show that fishermen who exhibit a higher propensity for cooperation in a laboratory public goods experiment, and those who show more patience in a laboratory time preference experiment, use fishing instruments which are less exploitative for the fishing grounds. We thus provide direct evidence that other-regarding preferences play an important role in crucial economic decisions in naturally occurring situations and establish other-regarding and time preferences as two distinct motivations for cooperation in the field to sustain CPRs.

The second essay, "*How can the Behavioralist prevail in the Market? Performance on Naturally Occurring Markets and Individual Preferences*", studies the direct link between individual preferences and performance on naturally occurring markets where asymmetric information and reputation play an important role. This essay shows that other-regarding preferences are important for economic outcomes in markets because they can restrain impatient sellers from yielding to their temptation to engage in uncooperative behavior that leads to instant gratification but hampers market performance in the long-run. I find that impatient sellers who lack other-regarding preferences misrepresent quality more and achieve considerably lower prices for goods of similar quality compared to both patient sellers who lack other-regarding preferences and other-regarding sellers (regardless of their time preferences).

The third essay, "*An Exploration of Third and Second Party Punishment in Ten Simple Games*", identifies the motives behind punishment from unaffected third parties and affected second parties using a within-subject design in ten simple games and a classification procedure.

We find that the most parsimonious model explaining the pattern of punishment includes inequity-averse and selfish subjects, and that this holds for both third and second parties. Our findings cast doubt on the idea that second and third parties punish in an impartial or normative manner.

The findings in the first essay corroborate the validity of social preference (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Falk and Fischbacher, 2006) and reciprocity theories (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Cox et al., 2007) for understanding persisting behavior in naturally occurring situations and show that there is no insurmountable gap between the laboratory and the field. With regard to economic policies, the findings suggest that one way to constrain overexploitation is by shifting the perception of the costs of current overexploitation into the present. Likewise, belief management and information policies which take the inherent bandwagon effects caused by preferences for conditional cooperation into account may be used for the management of CPRs.

The findings in the second essay point to an important role of other-regarding preferences for economic outcomes in naturally occurring markets and thus corroborate the validity of social preference and reciprocity theories for understanding market outcomes. I provide both theoretical and empirical evidence that other-regarding preferences can help impatient sellers from performance-damaging behavior.

The findings in the third essay have implications for the different social preference and reciprocity theories. Pure reciprocity models like Rabin (1993), Dufwenberg and Kirchsteiger (2004), and Cox et al. (2007) fail to account for third party punishment and therefore should be not applied alone to explain punishment. Norm approaches (e.g. López-Pérez, 2008) face an unanticipated problem, as there seems to be no way to explain the punishment patterns as a reaction to a prior deviation from any sensible norm of distributive justice. In contrast, social preference models like Fehr and Schmidt (1999) and Falk and Fischbacher (2006) fare much better in explaining third and second party punishment. We argue that economic policies should consider the endowment of third parties who serve in courts, as referees, or as arbitrators because it is likely that their judgment in these cases depends on their endowment, meaning that poorer parties sanction less impartially than richer parties.

This thesis proceeds as follows. Chapter 2 presents the first essay "*Cooperativeness and Impatience in the Tragedy of the Commons*". Thereafter, chapter 3 deals with the second essay "*How can the Behavioralist prevail in the Market? Performance on Naturally Occurring Markets and Individual Preferences*" and the last chapter 4 presents the third essay "*An Exploration of Third and Second Party Punishment in Ten Simple Games*".

Chapter 2

COOPERATIVENESS AND IMPATIENCE IN THE TRAGEDY OF THE COMMONS

Chapter Overview

This paper examines the role of other-regarding and time preferences in the exploitation of common pool resources (CPRs) by combining laboratory experiments with field data. We study fishermen whose main, and often only, source of income stems from the use of fishing grounds with open access. The exploitation of a CPR involves a negative interpersonal and inter-temporal externality because individuals who exploit the CPR reduce the current and the future yield for both others and themselves. Accordingly, economic theory predicts that more cooperative and less impatient individuals should be less likely to exploit the CPR. Our findings support this prediction because fishermen who exhibit a higher propensity for cooperation in a laboratory public goods experiment, and those who show more patience in a laboratory time preference experiment, use fishing instruments which are less exploitative of the fishing grounds. Thus, we provide direct evidence that other-regarding preferences play an important role in crucial economic decisions in naturally occurring situations and establish other-regarding and time preferences as two distinct motivations for cooperation in the field.

2.1 Introduction

According to economic theory there is little cooperation in sustaining common pool resources (CPR) where individual and collective interests are in conflict. The standard assumption of pure self-interest implies that natural resources like fishing grounds or rain forests are overexploited, and that we are often trapped in an inevitable process that ends in the "Tragedy of the Commons" (Hardin, 1968). An additional aggravating factor for resource conservation is the propensity to discount future outcomes. The more impatient resource users are, the more they exploit natural resources.[1] Interestingly, observations from the laboratory point a more optimistic picture regarding the occurrence of cooperation in social dilemmas. Considerable evidence now shows that some individuals cooperate voluntarily to sustain CPRs or public goods in the laboratory (Walker et al., 1990; Andreoni, 1995; Ledyard, 1995; Cardenas, 2000; Fehr and Gächter, 2000; Camerer, 2003; Casari and Plott, 2003; Croson, 2008; Charness and Villeval, forthcoming), suggesting that some individuals have other-regarding preferences (Andreoni, 1990; Rabin, 1993; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004; Sobel, 2005; Falk and Fischbacher, 2006; Segal and Sobel, 2007). There is also laboratory evidence that individuals differ with regard to their degree of impatience (Benzion et al., 1989; Ashraf et al. 2006).

In view of the key role that economic theory assigns to individuals' preferences in the exploitation of common natural resources, we study whether other-regarding and time preference measures predict fishermen's propensity to exploit a CPR that constitutes their main, and often only, source of income. This is a nontrivial task because it requires laboratory preference measures and field data from the same fishermen. The problem is that field measures of preferences are often confounded by all sorts of factors – such as reputational incentives, budget or information constraints – while for laboratory preference measures it cannot be taken for granted that they predict people's behavior outside the laboratory (Karlan, 2005; Levitt and List, 2007). However, if it can be shown that laboratory measures of other-regarding and time preferences are significantly predictive of fishermen's behavior in the field we can kill two birds with one stone. First, we provide direct evidence that other-regarding preferences play an important role for crucial economic decisions in naturally occurring situations. Second, we identify with other-regarding and time preferences two distinct motivations for cooperation in the field to sustain CPRs.

[1] Farzin (1984) shows that this statement holds as long as capital requirements for exploiting natural resources are low. If capital requirements are high, however, impatient individuals might shy away from investing in technology that facilitates the exploitation of natural resources and, thus, exploit natural resources less. In our setting, capital requirements are relatively low.

We achieve these goals with the help of two unique data sets that both combine individual laboratory behavior with the individual decisions of fishermen pertaining to the use of certain fishing instruments. Our study takes place in Brazil and involves fishermen who live off shrimp or fish caught from a lake. As there is free access to the lake, they face a CPR dilemma in their daily lives. There is suggestive evidence that fishermen in this setting differ in their level of cooperation to sustain fishing grounds, i.e. they use different fishing instruments where the proportion of the catch consisting of small shrimp/fish which have not yet reached fertility varies (Cavalcanti, 2003). We have data from fishermen who catch shrimp (collected 2008) and fish (collected 2006) and both data sets include information about the harmfulness of their fishing instruments as well as their decisions in two laboratory experiments: a public goods experiment where free-riding is the dominant strategy, and a time preference experiment. If laboratory preference measures capture relevant aspects of preferences in the field, economic theories of other-regarding preferences predict that individuals who exhibit a higher propensity to cooperate in the laboratory public goods experiment (i.e. who demonstrate other-regarding preferences), and those who show more patience in the time preference experiment should use fishing instruments which exploit the CPR less for the following reasons: (i) a higher current exploitation reduces the current yield of other fishermen. Thus, ceteris paribus, other-regarding fishermen will impose fewer current negative externalities on others; and (ii) a higher current exploitation (in terms of small shrimp/ fish that have not yet reached fertility) also reduces the future yield for both others and themselves. Therefore, more other-regarding and less impatient individuals will impose fewer future negative externalities on others and themselves.[2]

Our data supports these predictions and shows that in both data sets laboratory other-regarding and time preference measures are important predictors of individual behavior in real world CPRs. Fishermen who are more cooperative in the public goods experiments and those who are patient in the time preference experiments, use fishing instruments which are less exploitative of the fishing grounds. More cooperative and patient shrimp fishermen use shrimp traps with bigger holes where small and infertile shrimp can escape (see Figure a in the appendix) and more cooperative and patient fishermen who catch fish use fishnets with larger mesh sizes in which only bigger fish are caught (see Figures b and c in the appendix). Note that the behavior in the laboratory public goods and time preference experiments is predicted to be independent because in the public goods experiment time preferences can play no role. We find indeed that fishermen who are impatient in the time preference experiments are neither more

[2] In principle, it could be that other-regarding individuals will only impose fewer future negative externalities on others if they are patient, implying that there is an interaction effect between other-regarding preferences and patience. We checked for this possibility and do not find a significant interaction effect of other-regarding preferences and patience for CPR exploitation.

nor less likely to cooperate in the public goods experiments. Therefore, our study identifies that *both* other-regarding and time preferences independently influence cooperation in a naturally occurring field situation. In addition, we find that fishermen exploit fishing grounds less, (i) the less they believe that other fishermen exploit the fishing grounds and (ii) the higher they perceive the risk of the depletion of the fishing grounds.

While field behavior is sometimes explainable by other-regarding preferences, there is no direct evidence that other-regarding preferences actually play an important role for crucial economic decisions in naturally occurring situations. The existing field observations (e.g. Feeny et al., 1990; Sneath, 1998; Ostrom, 1999; Ostrom and Nagendra, 2006) cannot exclude the possibility that cooperation behavior in the field is driven by social pressure or reputation effects (Kandori, 1991; Bandeira et al., 2005; Mas and Moretti, forthcoming). Bandeira et al. (2005) find that fruit-pickers work less if their effort has negative externalities on their co-workers but only on fruit fields where they can be monitored by their co-workers. Similarly, Mas and Moretti (forthcoming) find that cashiers work faster if they can be observed by a harder-working colleague but not if the harder-working colleague cannot observe their work speed. In addition, a field experiment by Landry et al. (2006) suggests that charitable donations are not necessarily a consequence of altruism or other-regarding preferences but are often motivated by status concerns or social pressure.

Falk (2007) observes that charitable donations can be increased when a gift is provided in a solicitation letter, indicating that there can be gift-exchange in the field. It remains unclear, however, whether this finding results from a unique and unrepeatable reaction to a positive surprise and whether gift-exchange is present in more competitive environments. In this regard, the study by Gneezy and List (2006) suggests that the positive effect of a gift vanishes quickly and the study by List (2006) observes that other-regarding behavior is not present in a competitive environment. One explanation to reconcile the previous findings is provided by Plott (1996) who proposes the "Discovered Preference Hypothesis". In his view, people may be uncertain which behavior is in their best interest when they face new decisions and unfamiliar environments, and, therefore behave differently than when they make repeated decisions (for empirical evidence see Hoeffler and Ariely, 1999). An alternative explanation is that people have "constructed" and unstable preferences which depend on task and context, for instance people may be other-regarding in the context of donations but not in a different context (e.g. Slovic, 1995; Kahnemann, 1996).

Our study differs from earlier studies in that we combine both laboratory and field data from the *same* individuals to *directly* study the extent to which individual laboratory preference measures predict *persisting* individual field behavior in a competitive environment. This is an important distinction because without laboratory preference measures from the same individuals there is no direct evidence that other-regarding behavior in the field comes from other-regarding preferences and that individuals have stable other-regarding preferences.

The first existing study combining laboratory cooperation behavior with field data from the same individuals is Karlan (2005).[3] He conducted economic experiments with borrowers in a Peruvian microcredit program and reports that the behavior in a trust game predicts loan repayment. Individuals who transfer less money back to their trustor in a trust game are also more likely to drop out of the program due to defaults on their loans. This study shows that behavior in laboratory experiments can predict field behavior. However, since in his trust game trustors knew the identity of their trustees any back-transfer can be reconciled with selfishness in the presence of reputation effects and the role of other-regarding preferences therefore remains unclear in this study.[4] Moreover, it is not clear whether defaults actually measure other-regarding behavior because loan repayment is not completely under the control of the borrower. For instance, it is possible that some borrowers were not able to repay their loans due to unfavorable situations which can happen to both selfish and cooperative individuals. In contrast, our experiments are played anonymously and our field measures of fishing instruments capture other-regarding behavior in a clear manner: Fishermen choose (and build) their fishing instruments and, therefore, the level of CPR exploitation lies in their own hands.

Our findings help to assess the importance of other-regarding preferences as well as the scope of social preference and reciprocity theories in naturally occurring situations. We show that there is no insurmountable gap between the laboratory and the field and that the meaningful relations between laboratory and field behavior corroborate economic models that assume stable preferences. In addition, our evidence that both other-regarding preferences and impatience are important for the exploitation of CPRs suggests ways in which the overexploitation of CPRs may be reduced. Economic policies, for example, which shift the perception of the costs of current overexploitation into the present are likely to constrain overexploitation. Likewise, belief management and information policies which take into

[3] Benz and Meier (forthcoming) find a positive correlation between students' donations in a laboratory experiment and their charitable donations to a student fund. Similarly, Laury and Taylor (2008) observe that the behavior in a public goods experiment can be informative for understanding donations to a non-profit environmental organization. They observe behavior during a laboratory experiment in which students could also donate money to an environmental organization. This contrasts with our study which predicts the fishermen's cooperation behavior in the field in their professional activity – which constitutes their main source of income – with laboratory measures of other-regarding and time preferences.

[4] A similar argument also applies to the public goods game in Karlan (2005) which was also not played anonymously. In addition, it was a step level public good with many Nash equilibria, i.e., even purely selfish players had an incentive to contribute if they believed they were the pivotal players.

account the inherent bandwagon effects caused by preferences for conditional cooperation may be used for the management of CPRs.

The paper proceeds as follows. Section 2.2 presents the field setting and the field data. Section 2.3 presents the laboratory experiments. Section 2.4 links the behavior of shrimp fishermen in the laboratory experiments with their field data. Section 2.5 provides a robustness check for our results in the previous section by linking the behavior of a different set of fishermen in the laboratory experiments with their field behavior. Section 2.6 concludes.

2.2 Field Setting and the Data

2.2.1 Field Setting

Our study took place at a lake in northeastern Brazil. Several rural fishing villages[5] are situated around this lake where fishing is the main and often the only possible profession. Most fishermen catch shrimp and fish on their own, sell their catch at fish markets and thus provide their family with nutrition and income. There is free access to the fishing grounds (shrimp and fishing grounds), and a fisherman's capital requirements are rather low. For catching shrimp, fishermen only need a small boat and shrimp traps which they manufacture from used PET bottles.[6] While fishing, fishermen are typically scattered over the lake and fish at their preferred, sometimes remote spot(s). Other fishermen usually respect these spots, i.e. most fishermen do not fish at or close to another fisherman's spot. Their respect for others' fishing spots means that the fishing ground at this lake is not a pure CPR, but shares some features of a private property. The fishermen are aware of the fact that overfishing has negative externalities on others but the private aspect of their fishing spots also means that private investments like refraining from catching small shrimp or fish affect their own chance to catch these same shrimp or fish at a larger size at the same spot in the future.

There are no legal constraints concerning the studied fishing instruments and there are no legal regulations concerning the catching of shrimp.[7] In recent years, many fishermen have

[5] We use the term "villages" for reasons of simplicity. In this field setting, these are sometimes not villages in the ordinary sense, but rather community agglomerations where the borders between the neighboring community agglomerations or villages are unclear.

[6] For catching fish, fishermen typically use a fishnet. The costs for a fishnet can be normally paid with the income generated from one week's catch.

[7] There is just one legal regulation concerning the catching of fish which is the prohibition of catching small fish (below 20–30 centimeters, depending on fish type). This regulation is, however, not enforced.

complained about decreasing catch rates, which they mostly blame on the overexploitation of the shrimp and fish resources, that is, the catching of large amounts of small and infertile shrimp and fish (Cavalcanti, 2003). A research project with fishermen revealed their strong concern about the excessive exploitation of shrimp resources in this field setting (Cavalcanti et al., 2008). Governmental and local university institutions have taken note of the severity of the situation and first steps have been initiated to help sustain the fishing grounds. A management council has been introduced which is examining the current fishing situation.

2.2.2 Field Data

In the following sections, we report the data from the fishermen who catch shrimp. The data for the fishermen who catch fish is presented briefly in the robustness check in section 2.5. A more detailed version can be found in our working paper (Fehr and Leibbrandt, 2008). We investigated the fishing instruments from 114 fishermen from March - May 2008 and have data from their behavior in laboratory experiments in April 2008.[8] In addition, we conducted surveys in April 2008 where we collected socio-economic data and we measured cognitive skills in September 2008.

During the laboratory experiments and surveys, participants received a code to ensure anonymity and were free to leave at any point in time. The majority of our participants were male (73 percent), experienced and full-time fishermen (average years in profession was 17.4 years and average weekly work time was 21.1 hours, not including the time spent for preparing and selling the shrimp) who generated their income mainly from fishing. 57 percent derived their income exclusively from fishing, 29.5 percent derived a small additional income from selling agricultural products and 16.1 percent also had a part-time job. On average, fishermen reported to have a monthly income of 302.4 Reais from all activities (variable: *income*).[9] There were few differences in monetary wealth among our participants.[10] Fishermen from this region were in general poor and owned few valuable things besides their cottage. Fishermen spent on average 3.4 years in school and lived in a household with 5.5 persons (variables: *schooling and*

[8] We conducted eight experimental sessions in eight villages and the average session size was 26. Most of the experiments were run in local school buildings. We conducted experiments with in a total of 216 participants. In this paper, we report the behavior in a public goods and time preference experiment from participants where we know the average size of the shrimp traps (N=114). The full data set is available upon request.

[9] The Brazilian currency is called Real (singular) or Reais (plural). 1 Real equaled US $ 0.47 in September 2008, 302.4 Reais = US $ 142.1.

[10] We asked participants how wealthy/poor they are compared to other fishermen. 92 percent responded that they are not considerably more/less wealthy than other fishermen.

household size). 95.4 percent of the fishermen used only modified PET bottles with small holes to catch shrimp.[11]

We investigate the cooperation behavior of fishermen in sustaining fishing grounds by the fishing instruments they use to catch shrimp.[12] Investigating the fishing instruments has several advantages compared to other indices such as observing the catch quantity (or also the catch composition). First, by using fishing instruments as measures we have a robust picture of CPR exploitation and not just the snapshot that we would get when looking at catch rates/catch compositions for a limited period (fishermen may just have had a lucky/bad period and caught more/less than normally).[13] Second, the intention to exploit CPRs by using a fishing instrument with certain characteristics is obvious whereas deducing such an intention from the catch quantity is problematic (catching a large quantity of shrimp for instance, can be perceived as a skill and not necessarily as an uncooperative act among fishermen).

Fishermen manufacture their shrimp traps from used PET bottles and make many holes in these traps, i.e. they decide themselves how big the holes will be. The smaller the holes in the traps, the more small and infertile shrimp are caught in the trap.[14] Note that very small variations in the millimeter domain make a difference as to which shrimp are caught.[15] From March - May 2008 we collected one to two bottles from each of 114 fishermen and measured five to ten holes in each bottle of the 0.1 centimeter level, to construct our variable *hole size* which captures the average size of the holes in shrimp traps from one fisherman. Figure 1 shows the distribution of sizes fishermen use in our sample. The average hole size is 0.448 centimeters. Half of our participants use shrimp traps with a hole size between 0.367 and 0.500 centimeters, while 90 percent of the average hole size lies between 0.320 and 0.580 centimeters.

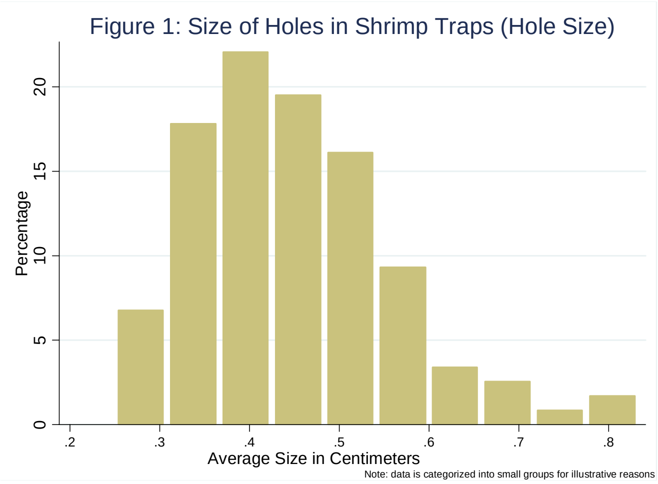
[11] The remaining 4.6 percent use two different shrimp traps at the same time (PET bottles and "bamboo baskets").

[12] In an earlier version of this paper, we also looked at the composition of the shrimp catch and the relationship to laboratory measures of cooperativeness and impatience. The findings were in line with the results presented here. For further information consult our working paper (Fehr and Leibbrandt, 2008).

[13] Of course, one could also collect data on catch quantity/composition over an extended period. In practice, such data is typically not available, and, moreover it would need a huge effort (and therefore be inefficient) to collect such data.

[14] The hole size is positively related to self-reported catch size. We asked fishermen to estimate how many liter shrimp they catch in general during a good week and find that the larger the hole size the less many of liter fishermen report catching (Spearman Rank Correlation, r =-0.246, p=0.0077). Note that our catch size measure is self-reported, imprecise and also inappropriate to capture CPR exploitation. This contrasts with our hole size measure that is not self-reported, very precise and also appropriate (because it captures the extent to which fishermen catch very small shrimp); and therefore is better suited to measure CPR exploitation.

[15] The average size of caught shrimp is usually between two and three centimeters.



We also collected survey data in 2008 regarding fishermen's perception and belief about current CPR exploitation in their setting. The variable *field perception* measures how fishermen perceive the risk that the shrimp population will be depleted in the near future due to the use of PET bottles per se. The variable *field belief* measures the fishermen's belief about the fraction of the other fishermen's shrimp catch that is below two centimeters. Moreover, we measure how central fishermen live by asking how many people live in their close surroundings (variable: *centrality*, average: 24.5). In addition, we have data on the cognitive ability of shrimp fishermen. We measured cognitive ability by giving fishermen three questions of the form: Which option is better for you, selling 75 liter shrimp for 1.2 Reais or 130 liter for 0.75 Reais. Fishermen had to respond quickly and were rewarded with 1 Real if they answered all three questions correctly. Overall, 18.7 percent answered all three questions correctly (two correct answers: 52.7 percent, one correct answer: 26 percent, zero correct answers: 2.7 percent).

2.3 The Laboratory Experiments

2.3.1 The Public Goods Experiment

Shrimp fishermen took part in an anonymous laboratory public goods experiment (PGE) with comparatively high monetary stakes (participants earned approximately 1.8 times their available daily income during the experimental sessions). The participants were divided into groups of three and played this experiment for one period. The payoff function was: $\Pi_i(x_i, x_j) = (10 - x_i) + 0.5 \times \left(\sum_{j \neq i} x_j + x_i \right)$. Each fisherman had to decide how many out of ten monetary units (MUs) he contributed ($x_i \in \{0,10\}$) to a public good.[16] For each unit he contributed, he increased the monetary payoff of each of his group members *j* by 0.5 MUs, but at the same time, his own balance was reduced by 0.5 MUs. For each unit one of his group members decided to contribute, his own balance was increased by 0.5 MUs. Since the fishermen's net return from contributing was negative, selfish fishermen should never contribute. However, if all four individuals in the group decided not to contribute, each of them only earned 10 MUs (10 – 0 + 0) compared to 15 MUs (0 + 0.5 × 10 × 3) if all of them contributed all ten MUs. The experiment was neutrally framed. Fishermen decided how many of the ten MUs they sent to a group account or otherwise kept in their private account. They were given two envelopes, one containing ten MUs and one containing 0 MUs. The participants could transfer MUs and then put both envelopes in a box. As the decision was made, the experimenter turned his back so that the fisherman was sure that the experimenter did not know his decision. After participants made their decisions, they were asked about their expectations of others' contributions. If they guessed the contribution of another participant correctly, they could win five MUs additionally. All rules were explained individually to the fishermen. No fisherman was informed about the identity of his group members.

Most fishermen do not behave according to the standard assumption but contribute to the public good (only 15.8 percent do not contribute and 11.4 percent contribute one MU). 21.1 percent contribute five MUs and 18.4 percent contribute more than five MUs. Approximately half of the participants contribute no more than three MUs (58 out of 114) and we denote them *low cooperative.* The remaining 49.9 percent contribute at least four MUs and we denote them *highly cooperative*. In regression table a in the appendix, we observe that expectations about the contributions of the other group members are by far the most important variable for predicting the behavior in the PGE (t>5.09, p<0.0001); which supports earlier evidence that many individuals are conditionally cooperative (Fischbacher et al. 2001, Keser and van Winden, 2002; Frey and Meier, 2004; Shang and Croson, 2008). In addition, we find that several control variables are marginally significant. We observe that more experienced fishermen contribute more (t>1.97, p<0.052), that fishermen who work more hours per day contribute marginally less (t>1.83, p<0.071), that fishermen who live in larger households contribute less (t>1.95, p<0.055) and that fishermen who report a higher monthly income contribute more (t>1.88, p<0.064). We will control for these variables in the following regressions.

[16] One MU always equaled one Real if the PGE was selected for payment. Participants only knew which experiment was actually paid out after they had played all experiments. The instructions for the laboratory experiments are in the appendix section A.1.1.

2.3.2 The Time Preference Experiment

At the beginning of the experimental sessions in 2008, we implemented a time preference experiment (TPE) to obtain a measure for impatience. In this TPE, all fishermen had to indicate whether they preferred two pralines immediately or three pralines at the end of the experimental session when receiving their payment (on the same day). The pralines ("Sonho de Valsa") are very popular among the fishermen; the vast majority (97 percent) liked the pralines.

We have data from this experiment for 83 of our 114 shrimp fishermen. 61.45 percent are patient and prefer to wait until the end of the experimental session to get three pralines whereas the remaining 38.55 percent are impatient and prefer to receive two pralines immediately. In column 2 of regression table a, we observe that none of the control variables significantly predicts the decision in the TPE.

2.3.3 The Relation between Other-regarding Preferences and Impatience in the Laboratory

Other-regarding preferences and impatience might play an important role in explaining the individual degree of CPR exploitation in the field. The laboratory provides an opportunity for deriving distinct measures for both factors. We use our PGE to obtain an individual measure for cooperativeness and the TPE to obtain an individual measure for impatience. Because there are no inter-temporal spillovers in the PGE, measured impatience in the TPE should not predict cooperativeness in the laboratory. This prediction is confirmed by our data. Individuals who are impatient in the TPE are not more or less likely to contribute more MUs in the PGE (Fisher Exact Test, $p=0.574$). This is also true after controlling for covariates. In regression table a in the appendix in models 3 and 4, we observe that the impatience dummy is insignificant ($p>0.310$). Note that there is also no relation between the public goods and time preference experiment that we conducted in 2006 ($p>0.820$, see also Fehr and Leibbrandt, 2008).

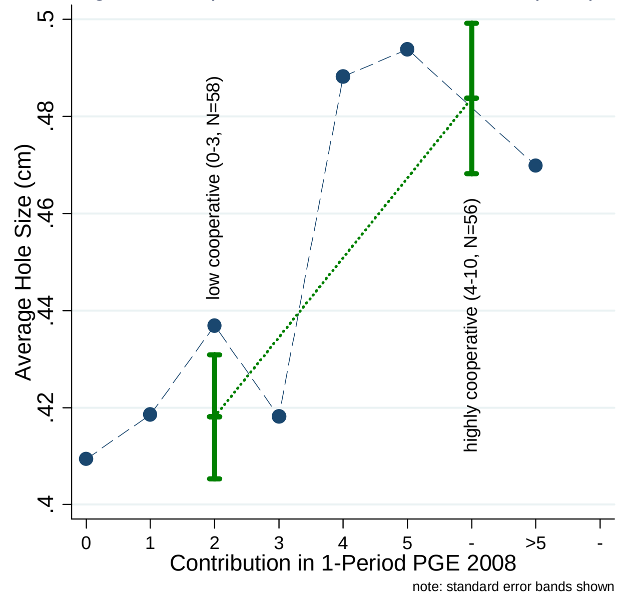
2.4 Predicting Cooperation in the Field in sustaining CPR with Laboratory Preference Measures

The heart of this paper uses our laboratory preference measures to predict individual levels of CPR exploitation. We use the public goods and time preference experiment from 2008 to predict the average *hole size* in the shrimp traps of fishermen. We hypothesize that fishermen who contribute more in the PGE use fishing instruments that are less exploitative of the fishing grounds, and that fishermen who are impatient in the TPE use fishing instruments that are more exploitative of the fishing grounds.

2.4.1 Other-regarding Preferences and Size of Holes in Shrimp Traps

Figure 2 provides a first insight into the relationship between contributions in the PGE and the average hole size in the shrimp traps. The *low cooperative* fishermen (who contribute less than four MUs; N=58) use on average a hole size of 0.418 centimeters whereas the *highly cooperative* fishermen who contribute at least four MUs (N=56), use holes that are more than 15 percent larger (0.484 centimeters). This difference is significant at any conventional level (t=-3.27, p=0.0014). The figure also shows the average hole size for fishermen who contribute one, two, three, four, five or six to ten MUs which is illustrated by the dots that are connected by a dashed line. It can be seen that the line is mostly increasing. The biggest difference exists between selfish fishermen who contribute zero (average is 0.409 centimeters, N=18) and fishermen who contribute five out of ten possible MUs (average is 0.494 centimeters, N=24). Fishermen who contribute between six and ten MUs use on average holes of 0.470 centimeters (N=21).

Figure 2: Cooperativeness and Holes in Shrimp Traps



In regression table 1, we investigate whether this effect is robust in different specifications. Model (1) shows that without using controls, each MU contributed is associated with a 0.0105 centimeters increase in average hole size (t=2.58, p=0.011) which means that a maximally cooperative fishermen uses on average 0.105 centimeter larger holes (i.e. approximately 25 percent larger holes). Model (3) shows the effect after adding several controls. The effect is even more pronounced and highly significant (t=2.85, p=0.005). In model (4), we additionally control for impatience which does not affect our previous findings, i.e. cooperativeness is still highly significant (t=2.48, p=0.016). In model (5), we introduce village fixed effects to account for potential regional differences. The effect of cooperativeness is slightly less pronounced, but still highly significant (t=2.15, p=0.036). In model (6), we control for cognitive skills which further reduces the sample size. Nevertheless, the effect of cooperativeness is strong and robust (t=2.05, p=0.046) and we can see that cognitive skills are completely uninformative for hole size (t=-0.24, p=0.815). Also when looking at the pure correlation between cognitive skills and hole size, we find no relationship (Spearman Rank Correlation, r=-0.01, p=0.926), i.e. cognitive skills as measured by calculation abilities are completely unrelated to hole size.

TABLE 1—DETERMINANTS OF SIZE OF HOLES IN SHRIMP TRAPS
(*OLS*)

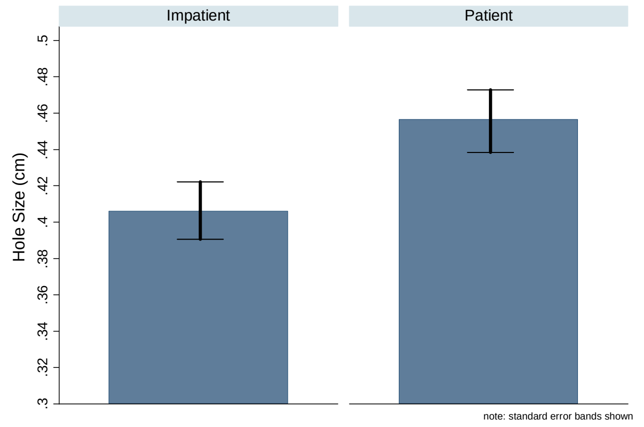
Dependent Variable	Average Size of Holes in Shrimp Trap in cm					
Model	(1)	(2)	(3)	(4)	(5)	(6)
Contribution in PGE	0.0105**		0.0106***	0.0104**	0.0088**	0.0093**
	(0.0041)		(0.0037)	(0.0042)	(0.0041)	(0.0046)
Impatience (Praline Dummy)		–0.0504**		–0.0544**	–0.0475**	–0.0566*
		(0.0230)		(0.0242)	(0.0229)	(0.0300)
Belief in PGE			0.0000	0.0011	0.0005	0.0080
			(0.0045)	(0.0051)	(0.0050)	(0.0067)
Impatience: likes praline?				–0.1658	–0.1352	
				(0.1621)	(0.1479)	
Age			0.0006	0.0013	0.0004	0.0022
			(0.0008)	(0.0009)	(0.0009)	(0.0014)
Gender (Male Dummy)			0.0927***	0.0568**	0.0415	0.0287
			(0.0227)	(0.0236)	(0.0325)	(0.0349)
Household Size			0.0070*	0.0059	0.0049	0.0066*
			(0.0036)	(0.0040)	(0.0041)	(0.0036)
Centrality			–0.0004	–0.0005	–0.0007*	–0.0004
			(0.0003)	(0.0004)	(0.0004)	(0.0003)
Schooling			–0.0063*	–0.0007	–0.0027	–0.0035
			(0.0035)	(0.0047)	(0.0046)	(0.0046)
Years in Profession			–0.0012	–0.0023**	–0.0021*	–0.0031**
			(0.0010)	(0.0011)	(0.0011)	(0.0014)
Field Belief			–0.0139	–0.0278**	–0.0233	–0.0412**
			(0.0110)	(0.0138)	(0.0148)	(0.0170)
Field Perception			0.0307***	0.0199**	0.0167*	0.0081
			(0.0095)	(0.0097)	(0.0098)	(0.0136)
Hours Fishing			0.0058	0.0086	0.0040	0.0065
			(0.0067)	(0.0061)	(0.0063)	(0.0087)
Quantity of Shrimp Traps			–0.0001**	–0.0001	–0.0000	–0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
Income			–0.0000	–0.0000	–0.0000	–0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
Cognitive Skills						–0.0035
						(0.0150)
Constant	0.4122***	0.4061***	0.2102**	0.4387**	0.4743**	0.4008***
	(0.0174)	(0.0164)	(0.0813)	(0.1860)	(0.1909)	(0.1126)
Village Fixed Effects?	no	no	no	no	yes	yes
Observations	114	83	113	83	83	66
R^2	0.064	0.051	0.322	0.377	0.450	0.494

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust Standard Errors in Parentheses.

2.4.2 Impatience and Size of Holes in Shrimp Traps

Figure 3 illustrates that fishermen who are impatient in the TPE and prefer two pralines immediately over three pralines at the end of the experimental session use smaller holes in their shrimp traps than patient fishermen. The average hole size for impatient fishermen is 0.406 centimeters whereas patient fishermen who prefer three pralines at the end of the experimental session use holes that are on average 0.457 centimeters, i.e. approximately 12 percent larger (t=-2.09, p=0.0396, two-sided).

Figure 3: Impatience and Holes in Shrimp Traps



This difference in the size of the holes in the shrimp traps is robust as we can see in the models 2, 4-6 in regression table 2. In model (2), we can see the pure effect after controlling for whether the fishermen like the pralines. A patient fishermen has on average 0.05 centimeter larger holes (t=2.19, p=0.031). In model (4), we use our standard control variables which does not change the significance of the impatience dummy (coefficient=0.054 centimeter, t=2.25, p=0.028). Impatience remains also significant at the 5 percent level after controlling for village fixed effects (model 5). In model (6) with the smaller sample, the impatience dummy is significant at p=0.065. None of the other variables besides cooperativeness and impatience are significantly predictive of hole size in all models. There are, however, some variables which seem to play an important role for hole size. Field perception predicts hole size in all models besides model 6, which may be due to the smaller sample size in this model. Fishermen who perceive a higher risk that the shrimp population will be depleted in the near future tend to use larger holes in their shrimp traps. The variable *field belief* is significant in models 4 and 6, and marginally insignificant in models 3 and 5. Fishermen who believe that other fishermen catch a larger fraction of small shrimp use smaller holes, suggesting that fishermen are also conditionally (un-) cooperative in the field. Male fishermen tend to use larger holes but this effect becomes insignificant after controlling for village fixed effects. Experience as a fisherman is negatively related to hole size, but this effect seems to be partly explainable by multicollinearity between experience and age. When we drop age in the regression, the effect of experience gets weaker and becomes insignificant or is only marginally significant. Household size is marginally positively related to hole size, suggesting that fishermen with more children are more concerned about sustainable fishing. The quantity of shrimp traps is significantly predictive of hole size in model 3 which shows that fishermen who use larger holes also use less shrimp traps.

2.5 Robustness Check

In 2006, we investigated the cooperation behavior of fishermen in sustaining fishing grounds (Fehr and Leibbrandt, 2008). Most fishermen use the fishnet as the fishing instrument as their fishing instrument of choice to catch fish. The fishnets differ according to their *mesh size* and the smaller the mesh size of the fishnet, the more infertile fish are caught in the fishnet.[17] We hypothesized that more cooperative fishermen use fishnets with larger mesh sizes and impatient fishermen use smaller mesh sizes.

2.5.1 The 2006 Field and Laboratory Data

We collected data on the mesh sizes of the fishnets from two sources: survey responses in 2006 and field observations in 2008. While re-visiting the fishermen in 2008, we investigated the fishnets of approximately every third fisherman who participated in 2006 and who used a fishnet (35 out of 121).[18] We find that the two year old survey responses are very much in line with the recent field data (Spearman Rank Correlation, $r=0.70$; $p<0.0001$). Figure d in the

[17] Note that fishnets differ according to their mesh size. However, the price of the fishnet is in this setting independent of the mesh size. Fishnets with smaller mesh sizes are not more expensive than fishnets with bigger mesh sizes. In case fishermen possessed more than one fishnet, the variable *mesh size* specifies the mesh size of the fishnet that is used most frequently. Using a fishnet with a larger mesh size can lead to an income reduction of approximately eight percent. We asked fishermen to fill out a daily report for several weeks where they reported which mesh size they used, the amount of fish they caught, and the weight in kilograms of fish they caught. Nine fishermen reported frequently using two different mesh sizes. When they use the smaller mesh size, they report catching a 21.5 percent larger number of fish per hour ($p < 0.01$) and 16 percent more kilograms of fish per hour ($p = 0.07$). If we assume that the additional fish caught with a small mesh size are all small fish that are sold at a 50 percent lower price (e.g., instead of a normalized price of 1 for larger fish the small fish are sold at a price of 0.5), the fishermen who use a small mesh size earn roughly 8 percent (0.16 x 0.5) more income per hour. A 50 percent lower price for small fish is a realistic assumption.

[18] Typically, we went to their houses or to the lake and asked them to show us their fishnet.

appendix illustrates the relation between the survey data and the field observations. Almost two-thirds of the fishermen use a mesh size that is smaller than five centimeters, with a mean of 4.42 centimeters. The most frequently used mesh sizes are 3.5, 5, and 6 centimeters. Figure e in the appendix shows in detail the mesh sizes fishermen use in our sample.

The fishermen also took part in an anonymously played laboratory public goods and time preference experiment (PGE 06 and TPE 06).[19] In the PGE 06, fishermen had to decide how many out of seven MUs they contributed to a public good in each of five different periods. The group size was four and stable during all periods. Because the fishermen's net return from contributing was negative, never contributing was always in their material interest if the selfishness and rationality of all individuals was common knowledge. The more cooperative the fishermen are, however, the more they should contribute. 87 percent contributed in the first period and almost half of the fishermen contributed between three and five units. In the remaining four periods, contributions declined continuously.

In the TPE 06, fishermen had to indicate whether they prefer one bottle of mineral water immediately or two bottles the next day. If they preferred the good immediately, the fishermen received the good immediately after the experiment. If they preferred two units of the good the next day, we distributed vouchers with which they could collect their good the next day at the village leader's house. The village leader was elected by the residents and is usually considered an extraordinarily trustworthy person.[20] We observe that 59.6 percent were impatient and preferred one bottle of mineral water immediately. As in our PGE 08 and TPE 08, we find no significant relationship between cooperativeness and impatience in the PGE 06 and TPE 06.

2.5.2 Cooperativeness, Impatience, and Fishnet Mesh Size

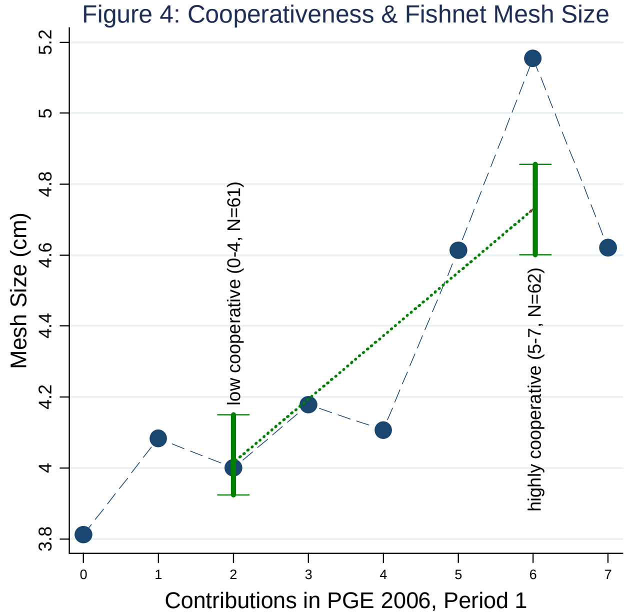
Figure 4 illustrates that there is a positive relationship between behavior in the PGE 06 and the mesh size. For instance, if we divide the fishermen into two equally sized groups according to their contributions in the first period of the PGE, we observe that those who are less cooperative and who contribute less than five MUs (N=61) use on average a mesh size of 4.03

[19] The instructions are in appendix A.1.2.

[20] In practice, the participating fishermen were not concerned about not receiving their good (the next day). Nevertheless, before individuals made their choice in the TPE, we assured them that they would receive their good. We also asked many participants if they were concerned about not receiving their good – which was not the case. The experimenters were also not strangers to the participants. At least one of the experimenters was known to the community leader before the experiments (who had her/his contact details) and to other fishermen. Participating fishermen also had the possibility to see how the experimenters gave the community leaders the mineral water for handing out the next day. After the experiments we got in contact with the community leader to ask whether all participants collected their goods – which was the case. When we re-visited the participants, none complained about not receiving her/his good.

centimeters (the mode is 3.5 centimeters, N=25) whereas those who are more cooperative and contribute at least five MUs (N=62) use on average a mesh size of 4.73 centimeters (the modes are 5 and 6 centimeters, each N=15). This difference is significant at any conventional level (T-test, t=-3.891, p=0.0002) and substantial.[21] The figure shows also the average mesh size for fishermen who contribute one, two …, seven MUs, which is illustrated by the dots that are connected by a dashed line. We can see that the line is increasing to the right and that the biggest difference exists between selfish fishermen who contribute zero (the average is 3.81 centimeters) and fishermen who contribute six out of seven possible MUs (the average is 5.15 centimeters).

Figure 4: Cooperativeness & Fishnet Mesh Size



In regression table 2 we use three OLS models to predict the mesh size. In model (1), we use the behavior in our two laboratory experiments as independent variables and control for several socio-economic and fishing related measures. In model (2), we control in addition for village fixed effects. We find that in both models, contributions in the first period of the PGE 06 are positively related to the mesh size of the fishnet. Each MU contributed in the first period of the PGE is associated with a 0.145/0.133 centimeter larger mesh size ($t > 2.72$, $p < 0.008$). In model (3), we use only the small sub-sample of 35 fishermen where we observed the fishnets and we control for the significant covariates from models 2 and 3. We still find a marginally significant relationship for cooperativeness (t=1.80, p=0.083). We also find a negative effect

[21] While it is difficult to find an exact correlation between a difference of one centimeter in mesh size with respect to the size of fish ultimately caught (since this does depend on the fish type), the fishermen estimate this to be approximately 3–7 centimeters. Note as a rough reference point that small fish that are below the legal minimum size (20–30 centimeters depending on fish type) are frequently caught in fishnets with mesh sizes that are smaller than five centimeters.

with regards to impatience in models 1 and 2. Fishermen who are impatient and prefer one bottle of mineral water immediately use 0.345/0.366 centimeter smaller mesh sizes ($t < 1.81$, $p < 0.073$) and this result even holds in the small sub-sample in model 3 ($t=1.86$, $p=0.074$).

In addition we find that the variable field perception plays an important role in all models. Fishermen who already perceive a relatively large mesh size as harmful tend to use larger mesh sizes. And we find that the variable field belief is significant in all models. It turns out that the more pessimistic a fisherman is about the exploitation level of the other fishermen, the more likely he is to use a small mesh size.

TABLE 2—DETERMINANTS OF FISHNET MESH SIZE
(*OLS*)

Dependent Variable	Frequently Used Mesh Size of Fishnet in cm		
Model	(1)	(2)	(3)
Contribution in First Period of PGE	0.145***	0.133***	0.171*
	(0.043)	(0.049)	(0.095)
Impatience (Mineral Water Dummy)	–0.345*	–0.366**	–0.641*
	(0.190)	(0.184)	(0.345)
Age	–0.011	-0.000	
	(0.012)	(0.013)	
Gender (Female Dummy)	–0.326	–0.732*	
	(0.335)	(0.374)	
Family Size	0.036	0.023	
	(0.033)	(0.037)	
Schooling	–0.012	–0.010	
	(0.032)	(0.033)	
Years in Profession	0.010	0.007	
	(0.010)	(0.010)	
Field Belief	0.178**	0.182**	0.503**
	(0.079)	(0.073)	(0.244)
Field Perception	0.160***	0.156***	0.352**
	(0.055)	(0.056)	(0.132)
Belief in first Period of PGE	-0.005	0.009	-0.144
	(0.053)	(0.059)	(0.141)
Hours Fishing	–0.006	–0.006	
	(0.007)	(0.006)	
Constant	3.022***	2.784***	2.023**
	(0.526)	(0.664)	(0.985)
Village Fixed Effects?	no	yes	no
Observations	121	121	35
R^2	0.256	0.334	0.227

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust Standard Errors in Parantheses.

2.6 Conclusion

In this paper, we show that other-regarding preferences play an important role for crucial economic decisions in naturally occurring situations. We find in two different data sets that, fishermen who behave more prosocially in a public goods experiment use fishing instruments that exploit the fishing grounds less. At the same time, we show that time preferences play an additional role for cooperation in the field to sustain CPRs and show in both data sets that fishermen who are impatient in a time preference experiment use fishing instruments that exploit the fishing grounds more. If only field or laboratory data were available, we could not uncover these relationships but the combination of the two allowed us to identify two different individual preferences that are both related to important cooperation decisions outside the laboratory.

Our results may have important implications for policymakers, managers, and social scientists. They provide empirical evidence that when designing policy measures it is useful to account for impatience as it is an obstacle for the implementation of resource preserving policies. Likewise, it may be useful to know about the conditional nature of fishermen's cooperativeness, i.e., their conditional willingness to cooperate in concrete situations even if cooperation goes against their immediate self-interest. Thaler and Benartzi (2004) designed the Smart commitment mechanism which helped impatient employees who lack self-control increase their future savings. We imagine a similar mechanism for promoting resource conservation which incorporates both the propensity to discount future outcomes as well as the propensity to cooperate voluntarily (if others cooperate as well). Individuals can be approached to commit in advance to change their behavior towards a more sustainable use of resources, but this commitment only becomes binding if a specified majority of the other resource users also commits. For example, the fishermen could commit (e.g., by signing a contract with an environmental agency) to exchange their fishnets with small mesh size to fishnets with bigger mesh size in the future if a specified majority of the other fishermen is also willing to commit to this policy. This proposal takes advantage of the conditional nature of fishermen's willingness to cooperate and, in addition, it reduces the perception of the cost of cooperation by shifting the exchange of the fishnets into the future. Thus, impatient individuals who lack self-control and conditionally cooperative individuals are more likely to commit to this policy than to an alternative policy that requires unconditional cooperation and imposes the cost of cooperation in the current period. Similar mechanisms can be applied to other settings as well. For example, policymakers could ask commuters in metropolitan areas to commit to buy a one-month ticket for public transportation for the following year as soon as a sufficient number of commuters is also willing to keep this commitment. Such a mechanism, which accounts for conditional cooperativeness and impatience, could discourage environmental unfriendly behavior and help sustaining natural resources.

2.A Appendix Tables and Figures

TABLE a—DETERMINANTS OF LABORATORY BEHAVIOR
(OLS, Probit)

Dependent Variable	Cooperativeness (OLS)	Patience (Probit)	Cooperativeness (OLS)	Cooperativeness (OLS)
Model	(1)	(2)	(3)	(4)
Patience Dummy (Praline)			0.386	0.539
			(0.587)	(0.527)
Belief in PGE	0.600***			0.564***
	(0.094)			(0.111)
Impatience: likes praline?		0.256		-0.046
		(0.391)		(1.527)
Age	-0.037	-0.005		-0.035
	(0.028)	(0.007)		(0.038)
Gender (Male Dummy)	0.204	0.040		0.019
	(0.412)	(0.138)		(0.544)
Household Size	-0.143**	-0.012		-0.173*
	(0.065)	(0.021)		(0.089)
Centrality	-0.010	0.002		-0.008
	(0.006)	(0.002)		(0.006)
Schooling	0.056	-0.023		0.139
	(0.113)	(0.024)		(0.146)
Years in Profession	0.050*	-0.002		0.066**
	(0.026)	(0.007)		(0.029)
Field Belief	-0.243	0.061		-0.206
	(0.279)	(0.071)		(0.272)
Field Perception	-0.100	0.026		-0.113
	(0.240)	(0.060)		(0.262)
Hours Fishing	-0.254*	-0.009		-0.349**
	(0.139)	(0.037)		(0.162)
Income	0.002**	0.000		0.002*
	(0.001)	(0.000)		(0.001)
Observations	113	83	83	83
R^2	0.412		0.005	0.398
Pseudo R^2		0.055		

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust Standard Errors in Parantheses.

Figure a: Shrimp Trap made of PET bottle

Note: Shrimp enter through big hole in front and can only escape through small holes at the bottom of the bottle.

Figure b: Fishnet from one Fisherman

Figure c: Measuring the Fishnet Mesh Size

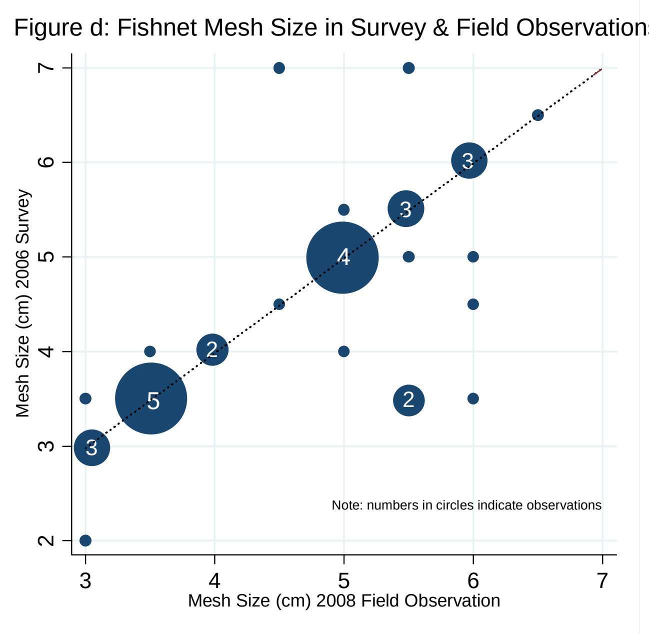


Figure d: Fishnet Mesh Size in Survey & Field Observations
Mesh Size (cm) 2006 Survey
Mesh Size (cm) 2008 Field Observation
Note: numbers in circles indicate observations
2
3
4
5
6
7
3
5
2
4
3
3
2

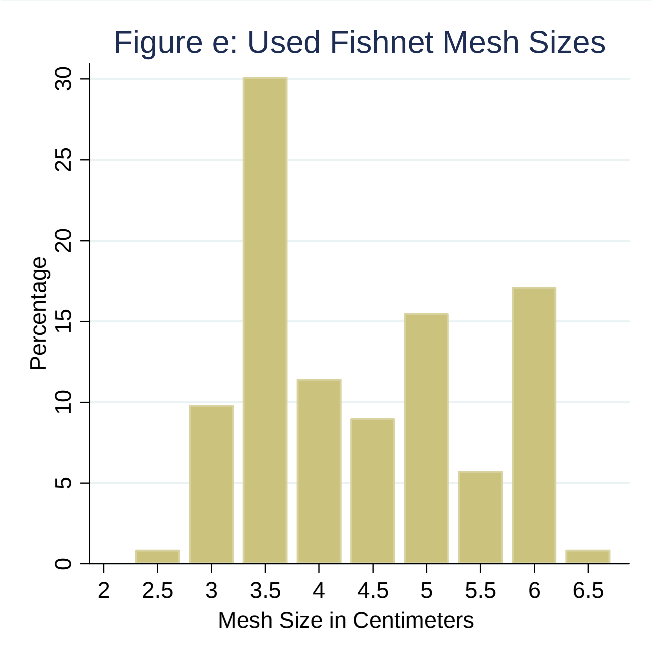


Figure e: Used Fishnet Mesh Sizes
Percentage
Mesh Size in Centimeters
0
5
10
15
20
25
30
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5

Chapter 3

HOW CAN THE BEHAVIORALIST PREVAIL IN THE MARKET? PERFORMANCE ON NATURALLY OCCURRING MARKETS AND INDIVIDUAL PREFERENCES

Chapter Overview

This paper experimentally studies the direct link between individual preferences and performance in naturally occurring markets where asymmetric information and reputation play an important role. Overall, I find that less cooperative sellers underperform compared to more cooperative sellers and achieve considerably lower prices for goods of similar quality. The overall finding is mediated by impatient less cooperative sellers who also misrepresent quality more than more cooperative sellers (patient and impatient) as well as patient less cooperative sellers. This shows that other-regarding preferences are important for economic outcomes because they can restrain impatient sellers from yielding to their temptation for instant gratification and thus prevent them from handicapping their market performance. Moreover, this paper provides new field evidence that the degree of risk-aversion is negatively related to market performance.

3.1 Introduction

Numerous laboratory experiments provide evidence that there are cooperative individuals (Roth, 1995; Camerer and Thaler, 1995; Camerer, 2003) and that the cooperativeness of some individuals results in more efficient outcomes in moral hazard contexts than the standard self-interest model assumes (Fehr et al, 1993; Brown et al., 2004; Fehr et al., 2007). These findings led to the formulation of theories of other-regarding preferences which provide several explanations for cooperativeness based on non-monetary motivations (Andreoni, 1990; Rabin, 1993; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004; Sobel, 2005 and 2007; Falk and Fischbacher, 2006; Cox et al., 2007; López-Pérez, 2008). However, it is still unclear how more cooperative individuals can prevail in markets and keep up with less cooperative individuals who are more likely to engage in uncooperative behavior whenever it is in their own monetary interest and therefore should tend to outperform more cooperative individuals.

In this paper, I investigate the behavior and performance of sellers with different individual preferences in naturally occurring markets. I present a simple model which predicts that other-regarding preferences are beneficial for market performance because they can restrain impatient sellers from yielding to their temptation to engage in uncooperative behavior that leads to instant gratification but hampers their market performance in the long-run. To verify this prediction and to identify the links between individual preferences, behavior, and performance in the field, the present study uses a unique data set that combines individual field and laboratory data from the same individuals. The studied individuals are fishermen who derive their income from selling shrimp at markets where asymmetric information about the quality of the shrimp – and reputation in general – play an important role. The data comprises information about the trades at these markets like achieved prices and offered quality, quality misrepresentations, and laboratory measures of the sellers' other-regarding, time and risk preferences.

I find that other-regarding preferences are indeed beneficial for impatient sellers and restrain them from misrepresenting quality and thus hampering their market performance. Impatient, less cooperative sellers (i.e. those who are self-interested or who have few pronounced other-regarding preferences) are more likely to misrepresent quality than either *patient*, less cooperative sellers or more cooperative sellers, regardless of their patience. Whereas, more cooperative sellers outperform less cooperative sellers overall and achieve significantly and considerably higher selling prices for shrimp of similar quality, the data reveals that time preferences mediate the performance difference between more and less cooperative sellers. Only impatient, less cooperative sellers underperform in comparison to patient, less cooperative and more cooperative sellers in general. In addition, I find that more risk-averse sellers underperform and achieve significantly lower selling prices for shrimp of similar quality compared to less risk-averse sellers. Interestingly, sellers who are less exposed to markets are endowed with similar preferences compared to sellers who are more exposed, suggesting that individual preferences affect market outcomes (and not vice versa).

This paper shows that other-regarding preferences play an important role in economic outcomes in naturally occurring markets and thus corroborate social preference and reciprocity theories. Importantly, it provides both theoretical and empirical evidence that more cooperative sellers can have competitive advantages compared to impatient, less cooperative sellers. Therefore, the data also refute the conventional wisdom claiming that markets make sellers selfish and challenges recent studies questioning the presence of other-regarding preferences in markets (e.g. List, 2006). To the best of my knowledge, this paper is the first to investigate the relationship between combinations of different individuals' preferences, behavior, and performance in naturally occurring markets.

Few studies consider the importance of individual preferences for performance in naturally occurring markets. The findings in my paper are supported by survey evidence drawn from large and representative sample sizes (Bonin et al., 2007; Dohmen et al., forthcoming). Bonin et al. (2007) find that individuals who report having a lower propensity to take risks are more likely to work in occupations with lower wages. Dohmen et al. (forthcoming) provide evidence on the relationship between inclinations to reciprocate and labor market outcomes. They find that individuals who are more likely to respond to questions like "If someone does me a favor, I am prepared to return it" with "applies to me perfectly" also report to working harder and having higher incomes.[22] In contrast, the present study uses behavioral preferences measures, actual instead of reported field behavior and outcomes, and, in addition, provides an explanation for the overall finding that more cooperative (i.e. more reciprocal) individuals can generate higher incomes.

The paper proceeds as follows. Section 3.2 presents the field setting and the field and laboratory data. In section 3.3, I present a model to derive predictions for the market behavior and performance of the different seller types. Sections 3.4 to 3.6 present the empirical data. Section 3.4 presents how other-regarding and time preferences are linked to quality misrepresentation. Section 3.5 investigates the relationships between other-regarding, time and

[22] A possible objection could be that it is likely that their survey questions not only capture inclinations to reciprocate, but other aspects, such as social desirability concerns. And it seems plausible that individuals who care more about social desirability (and therefore appear more positive reciprocal) are more successful at the labor market.

risk preferences and market performance. Section 3.6 presents whether different levels of market exposure affect individual preferences. Section 3.7 concludes.

3.2 Field Setting and the Data

3.2.1 Field Setting

The study took place in Brazil, using shrimp fishermen who live by selling their catch. In this setting, fishing is the main and often the only possible profession for generating income. There is free access to the fishing grounds and capital requirements for becoming a fisherman are low: only a small boat and shrimp traps that are made of used PET bottles are needed. Most fishermen sell their catch at markets and in this way cover their living expenses. In the study region, there is one big and several small markets where fishermen sell their shrimp once a week. Typically, they are able to sell their whole catch (which is often more than 100 liters) quickly; most of their catch is bought by one or several middlemen, who then resell it to bars and restaurants. Figure a in the appendix illustrates the shrimp market. Reputation plays a crucial role, as long-term trade relations between sellers and buyers are common.

The shrimp markets constitute an ideal field setting for studying market performance. *First*, the traded shrimp differ only in one important quality dimension, which is average shrimp size.[23] Bigger shrimp are tastier, and are sold for significantly higher prices per liter than smaller shrimp. *Second*, there is considerable variance in selling prices among the different sellers, even after controlling for shrimp size. This implies that some sellers achieve higher selling prices for similar shrimp, i.e. there are considerable differences in market performance.

There is asymmetric information about the precise size of the traded shrimp. The shrimp are presented in large piles and the bigger shrimp are often placed on top. A buyer cannot easily verify the average shrimp size from one seller on a particular day, unless he puts a large effort in scanning the piles, which the sellers typically do not appreciate.[24] Therefore, it can be

[23] Shrimp also differ in color because some fishermen color their shrimp red to make them look tastier. I control for color in some regressions but find no significant effect. Note that there are no visible differences in the shrimp type (there are two shrimp types which are extremely similar) and that freshness plays no important role because the shrimp are sold dried.

[24] Note that the average size of the shrimp differs from week to week, i.e., even if buyers measured the exact size in a certain week (after buying them), they would not know the exact size in a different week. Although sellers have some influence on the average shrimp size (as this depends on the holes in their shrimp traps), other factors not under control of the fishermen, like season or luck, also influence average shrimp size. I collected data on the average shrimp size over consecutive weeks from 24 sellers. Figure b in the appendix shows the average shrimp size of buyers in two different weeks. As expected there is a significant correlation of average shrimp size across weeks (Spearman Rank Correlation, $r=0.493$, $p=0.014$), but considerable variance as well.

beneficial for buyers to rely on the sellers' quality estimate if buyers believe the seller is trustworthy and reports quality truthfully.

3.2.2 Field Data

The field data comes from observations of actual trades on shrimp markets and from two different surveys. The field data about the trades was collected within a three week horizon in April-May 2008; the surveys were conducted individually in village meetings in April and August 2008. In total, there is field data from 99 fishermen which contain the selling price per liter shrimp, the average shrimp size, the quantity sold in liters, and the color of the shrimp. I use the selling price per liter shrimp as a measure for market performance because it is the most direct measure, and unlike other measures (e.g. income), it depends less on other imperfectly observable factors.[25] The data was typically collected at the moment the trades were concluded.[26] We also collected one liter of the sold shrimp from the bottom of the piles to identify the representative average size and color of the shrimp. We calculated the average size by averaging the size of 30 different shrimp to provide a reliable measure. We took the shrimp from the bottom because some fishermen put the bigger shrimp on top of the piles to misrepresent quality, i.e. to give the impression that the average shrimp size is bigger.

In addition, we collected two samples of shrimp, one from the bottom and one from the top of the pile, from 33 fishermen. It turns out that in 25 of the 33 samples, the average shrimp size is larger in the sample on the top. On average, the sample from the top contains shrimp that are 7.32 percent larger (one sample T-test, $t=3.12$, $p<0.004$ that mean equals zero). I use the percentage difference between the two samples as a measure for misrepresenting quality. Figure 1 presents the differences between top and bottom shrimp in more detail. On the right side of the vertical line, we can see the distributions of quality misrepresentations where sellers have larger shrimp on top of the piles.

---

[25] For instance, price per liter shrimp in our sample does not significantly depend on operating expenses (like hours fishing) or the quantity of shrimp sold, i.e. there are no quantity discounts. This means that sellers who achieve a higher price per liter shrimp can also achieve a higher income. Note, however, that sellers who achieve a higher price per liter shrimp do not necessarily have a higher income. For instance, these sellers might spend less time on the lake and catch fewer shrimp because they are income targeting (for evidence on income targeting see Fehr and Götte, 2008, for example).

[26] The experimenters stayed at the market and asked the fishermen in advance for their permission to do so. We ensured fishermen that nobody besides the experimenter will be able to retrace the individual terms of the trades.

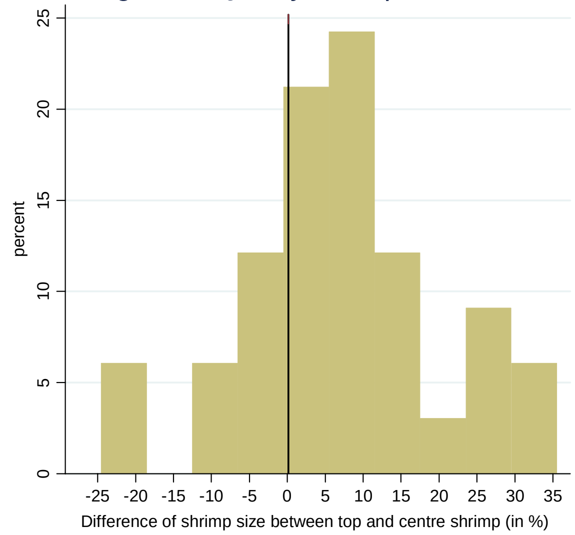


Figure 1: Quality Misrepresentation

Approximately 60 percent of the field data comes from the biggest market in the region, and the remaining observations come from three smaller markets. On average, the selling price per liter is 1.27 Reais[27]. The average shrimp size is 2.97 centimeters (50 percent lie between 2.74 and 3.23 centimeters). 44 out of 99 fishermen color their shrimp. The distributions for the selling prices and average shrimp sizes are presented in Figures c and d in the appendix. Figure e in the appendix illustrates that there is a significant positive relationship between selling prices and average shrimp sizes across all markets (Spearman Rank Correlation, r=0.320, p=0.001), and that the selling price increases on average by 0.386 Reais for each additional centimeter in average shrimp size. Table a in the appendix provides a summary statistic for the field data.

To test the existence of asymmetric information, buyers and sellers took part in a "guessing game", where the most accurate guess about the average shrimp size in a pile was rewarded with a high monetary reward.[28] It turned out that while buyers considerably overestimated the average size (by 0.437 centimeters) of shrimp of a seller from whom they frequently buy, sellers where much more accurate in guessing the average shrimp size of their own catch (+0.286 centimeters; Variance Comparison Test, p=0.013, two-sided), providing evidence for the persistence of asymmetric information in this setting.

In April 2008, a total of 216 fishermen took part in a survey, including the 99 fishermen for whom I collected the field data. The survey was conducted individually at meetings and

[27] 1 Real (pl. Reais) equaled approximately 0.60 US Dollars during the observation period.

[28] The guessing game took place twice at a big shrimp market. The best guess was rewarded with 100 respectively 50 Reais, which is approximately a one (half)-week income of a shrimp seller. In total, 25 sellers and 12 middlemen took part in the guessing game.

included questions about the socio-demographic characteristics of the fishermen. The typical fisherman is experienced and has already been fishing for 17.7 years. Fishermen sell their shrimp to friends, distributors, restaurants, or at markets. Approximately two-thirds (64.19 percent) always sell their shrimp at the market, 19.53 percent never sell their shrimp at the market, 9.3 percent rarely and 6.98 percent sometimes. There are two types of sellers: specialized shrimp sellers whose sales are limited to shrimp, which I denote by the dummy *shrimp seller*, i.e. they sell only shrimp (75 percent), and unspecialized shrimp sellers who switch between selling shrimp and fish (25 percent).[29]

In August 2008, 150 of the 216 fishermen who participated in the survey in April took part in an additional short survey where I collected data on cognitive skills. I measured cognitive skills by giving fishermen three questions of the form: Which option is better for you, selling 75 liter shrimp for 1.2 Reais per liter, or, 130 liter for 0.75 Reais per liter. Fishermen had to respond quickly and were rewarded with 1 Real if they answered all three questions correctly. 18.7 percent answered all three questions correctly (two correct answers: 52.7 percent, one correct answer: 26 percent, zero correct answers: 2.7 percent).

3.2.3 Laboratory Data

All 216 fishermen who participated in the survey took also part in April 2008 in a public goods and risk-aversion experiment (PGE and RAE) with considerable stakes (participants earned approximately 1.8 times their available daily income during the experimental sessions). The laboratory experiments did not take place at the markets, but were conducted individually during village meetings. In the PGE, the participants were divided in groups of three and played this experiment for one period.[30] The payoff function was:

$$\Pi_i\left(x_i, x_j\right) = (10 - x_i) + 0.5 \times \left(\sum_{j \neq i} x_j + x_i \right).$$

[29] Specialized shrimp sellers may be able to achieve higher selling prices. They can probably build up long-term relations more easily and know more about the dynamics of the shrimp markets because they usually go to the market every week to sell their shrimp, whereas the sellers who switch between catching fish and shrimp often interrupt their selling of shrimp for longer periods. I will take care of this concern and control for the seller type.

[30] Note that the group size in the PGE was four in one experimental session (N=16). The behavior in this session is very similar compared to all other sessions (average contribution in this session = 3.75, in all other sessions = 3.66; t=-0.12, p=0.90). Excluding this session from our further analysis would not lead to systematic changes.

Each fisherman had to decide how many out of ten monetary units (MUs) he contributes ($x_i \in \{0,10\}$) to a group account.[31] For each unit he contributed, he increased the monetary payoff of each of his group members *j* by 0.5 MUs, but at the same time, his own balance was reduced by 0.5 MUs. For each unit one of his group members decided to contribute, his own balance was increased by 0.5 MUs. Since the fishermen's net return from contributing was negative, selfish fishermen should never contribute. However, if all three individuals in the group decided not to contribute, each of them only earned 10 MUs (10 – 0 + 0), compared to 15 MUs (0 + 0.5 × 10 × 3) if all of them contributed all ten units. I denote the contribution decision as cooperativeness. The more fishermen contribute, the more cooperative they are, i.e. the more pronounced are their other-regarding preferences. After they made their contribution decision, they were asked about their expectations about the other group members' contributions. If they guessed them correctly, they could receive an additional five MUs. All rules were explained individually to the fishermen. No fisherman was informed about the identity of his group members.

Most fishermen contributed to the public good; only 16.2 percent did not contribute and 11.1 percent contributed only one MU, 19.44 percent five MUs and 14.8 percent more than five MUs. I discriminate the fishermen in two equally sized groups: *low cooperative* fishermen (= LC fishermen) who contributed between zero and three MUs (51 percent), and, *highly cooperative* fishermen (= HC fishermen) who contributed at least four MUs (49 percent).

In the RAE, fishermen had to decide how many out of ten MUs they invest in a lottery that provides a return of 2.5 times the risked MUs with a probability of 0.5 and a return of zero with probability 0.5. Since the expected return from the lottery was greater than one (2.5 × 0.5 = 1.25) risk-neutral and risk-seeking individuals should always invest all ten MUs in the lottery. Risk-averse participants, however, may invest less and very risk-averse participants should invest nothing or very few MUs.[32] I observe that 21.8 percent decided not to risk any MUs, 20.4 percent invested three, 23.6 percent invested five MUs and only 1.85 percent invested all ten MUs. I apply a median-split and discriminate two groups of fishermen: *highly risk-averse*

[31] The participants were given two envelopes, one containing ten MUs and one containing 0 MUs. They could then transfer MUs and put both envelopes in a box. During the time of the decision, the experimenter turned his back so that the fisherman was sure that the experimenter did not know his decision. One MU always equaled one Real if the PGE was selected for payment. Participants only knew after they had played all experiments which one was actually paid out. The instructions are in the appendix section A.1.1.

[32] The exact procedure of the RAE was as follows: Fishermen had to toss a coin and beforehand announce whether they would like to choose heads or tails and how many out the ten MUs they want to risk. If the announced side of the coin showed up, they would win 2.5 times the risked MUs, and if the other side showed up they would lose their risked MUs.

fishermen who invested less than three MUs (40 percent), and, *low risk-averse* fishermen who invested at least three MUs (60 percent).[33]

I also report data from a time preference experiment (TPE) conducted in 2006 where 51 of the 99 fishermen participated that I observed in 2008 at the markets. In the TPE, fishermen had to indicate whether they prefer one bottle of mineral water immediately or two bottles of mineral water the next day. If they preferred the mineral water immediately, they received one bottle directly after the experiment. If they preferred the two bottles the next day, I distributed vouchers with which they could collect the two bottles the next day in the village leader's house. The village leader was elected by the residents and is usually considered an extraordinarily trustworthy person.[34] Participants were able to witness the experimenters giving the leaders the mineral water which was to be distributed the next day.[35] The reason I chose mineral water was because it often measures immediate utility since individuals tend to consume it the moment they get it. This is not necessarily the case for money in the studied remote setting, where there are very few possibilities for spending money in the immediate vicinity. Overall, I observe that 58.6 percent were impatient and preferred one bottle of mineral water immediately.

Note that there is no significant relationship between the behavior in the PGE and the TPE. HC fishermen are not more or less likely to be impatient (Fisher's exact Test, p=1.00, two-sided). The behavior in the RAE is also unrelated to the behavior in the TPE (Fisher's exact Test, p=1.00) but significantly negatively related to the behavior in the PGE (Fisher's exact Test, p=0.018), i.e. LC fishermen are more likely to be highly risk-averse.

[33] I assign the fishermen who invested three MUs (20.4 percent) to the low risk-averse fishermen.

[34] The village leader is the elected representative of the village and frequently visits meetings to present the situation and needs in his village.

[35] I have no reason to believe that the participating fishermen were concerned about not receiving their good the next day. Nevertheless, before individuals made their choice in the TPE, we assured them that they will receive their good. We also asked many participants if they were concerned about not receiving their goods – which was not the case. The experimenters were also no strangers to the participants. At least one of the experimenters was known before the experiments to the community leader (which had her/his contact details) and to other fishermen. After the experiments, we contacted the community leader to ask if all participants collected their goods – which was the case. When we re-visited the participants, none complained about not receiving her/his good. The instructions are in the appendix section A.1.2.

3.3 Predictions for Quality Misrepresentation and Market Performance

In this section, I derive the predictions for the relationships between individual preferences and individual market performance. To do this, I present a simple two-period bargaining model which abstracts from the complex interactions between the different buyers and sellers in our competitive field setting. Nevertheless, the model can provide some intuition how the sellers' preferences can affect their behavior and outcomes. In this model, there is one buyer and one seller. The seller wants to sell one product to the buyer for a certain price. Because there is asymmetric information about the quality of the product, the seller can sell the product for a "fair price" $p = p_{FAIR}$ or an inappropriately high price $p > p_{FAIR}$. To achieve an inappropriately high price, the seller must misrepresent quality (for instance by putting the bigger shrimp on top of the pile). There are two different seller types: (a) selfish sellers and (b) cooperative sellers. Selfish sellers care only about the achieved price *p*, and their utility function in the two-period case is:

$$(1) \qquad U_S = p_1 + \delta_i \times p_2 \,,$$

with δ_i = discount rate of seller *i* ($0 \leq \delta_i < 1$), i.e. δ_i can be understood as time preferences with numbers closer to 0 indicating that *i* is impatient. In contrast, cooperative sellers derive disutility from selling a product for an inappropriately high price. One possibility for modeling this is to assume that cooperative sellers care about the difference between *p* and p_{FAIR}. Formally, I say that if $p > p_{FAIR}$, cooperative sellers derive disutility depending on λ_i ($\lambda_i > 0$) which is an individual parameter expressing the degree of disutility from selling a product for an inappropriately high price, and, the total disutility if $p > p_{FAIR}$ is $\lambda_i (p - p_{FAIR})$. Therefore, a cooperative seller's utility function in the two-period case is:

$$(2) \quad U_{COOP} = \lfloor p_1 - \lambda_i \max(p_1 - p_{1,FAIR}, 0) \rfloor + \delta_i \lfloor p_2 - \lambda_i \max(p_2 - p_{2,FAIR}, 0) \rfloor.$$

In order to predict the market performance of the different seller types, we need to know whether sellers will try to sell a product for $p > p_{FAIR}$ in *t=1*. This of course depends on whether buyers will find out in *t=2* about the inappropriately high price in *t=1* and how they react in this case. For simplicity, let us assume the following: (i) buyers know definitely in *t=2* whether an inappropriately high price was charged in *t=1*, i.e. the detection probability is set to one, (ii) if an inappropriately high price was charged in *t=1*, buyers are only willing to pay $p_{DETECTED}$ ($p_{DETECTED} < p_{FAIR}$) in *t=2*, i.e. they deduct a risk premium ($p_{FAIR} - p_{DETECTED}$) when buying a product from a buyer who cheated in the previous period.[36] Let us further assume that the risk premium is so high that a selfish seller with $\delta_i = 1$ will never misrepresent quality in *t=1* if the detection probability is set to 1, i.e. $p_{FAIR} - p_{DETECTED} > p - p_{FAIR}$. Accordingly, selfish sellers will charge $p > p_{FAIR}$ in *t=1* if

$$(3) \quad p + \delta_i \times p_{DETECTED} \geq p_{FAIR} + \delta_i \times p_{FAIR} \quad \Leftrightarrow \quad \delta_i \leq \frac{p - p_{FAIR}}{p_{FAIR} - p_{DETECTED}}.$$

As we observe, the decision for the selfish seller to charge an inappropriately high price in *t=1* depends on his δ_i, i.e. the more impatient the selfish seller is, the more likely he is to sell his product at an inappropriately high price in *t=1*.

PREDICTION 1: *(i) Impatient selfish sellers misrepresent quality more than patient selfish sellers. Therefore, the market performance of selfish sellers depends on their level of impatience; (ii) impatient selfish sellers achieve lower selling prices in all periods t>1 compared to patient selfish sellers.*

Let us now take a look at the behavior and performance of cooperative sellers whose disutility from selling a product for an inappropriately high price is so high that they would misrepresent quality in *t=1*. I denote these as *highly cooperative* sellers (= HC sellers). The following must apply for a HC seller:

$$(4) \quad p_{1,FAIR} + [\delta_i \times p_{2,FAIR}] > p_1 - \lambda_i(p_1 - p_{1,FAIR}) + \delta_i[p_2 - \lambda_i(p_2 - p_{2,FAIR})]$$

$$\Leftrightarrow \quad [1 + \delta_i] \times [p_{FAIR}] > [1 + \delta_i] \times [p - \lambda_i(p - p_{FAIR})]$$

$$\Leftrightarrow \quad p_{FAIR} > p - \lambda_i(p - p_{FAIR})$$

$$\Leftrightarrow \quad \lambda_i > 1$$

Thus, for a HC seller with $\underline{\lambda}_i = \lambda_i > 1$, the decision whether to sell a product for an inappropriately high price does not depend on his δ_i i.e. impatience plays no role for this seller type.

[36] Alternatively, we could assume that a buyer who cheated in *t=1* faces with a positive probability the risk that the trade relation is terminated in *t=2* and in this case cannot sell his product in *t=2* (see Brown et al. 2004).

PREDICTION 2: *(i) There are no differences in quality misrepresentation between impatient and patient highly cooperative sellers. Therefore, (ii) there are no differences in market performance between impatient and patient highly cooperative seller, both achieve similar selling prices in t>1.*

Note that predictions 2a and 2b are only true for HC sellers ($\lambda_i > 1$). For *low cooperative* sellers (= LC sellers) with $0 \leq \lambda_i > 1$, time preferences can affect quality misrepresentation. Thus, from now on I will always differentiate between LC sellers (which include the selfish sellers) and HC sellers.

Therefore, if we do not control for impatience we should find that,

PREDICTION 3: *overall, given that prediction 1 and 2 are true, HC sellers outperform LC sellers because one fraction of the LC sellers –the impatient– underperforms compared to HC sellers.*[37]

Risk preferences can also play an important role for market performance in our setting because a seller has to propose or accept a price without knowing a potential buyer's maximal acceptable price. Thus, a more risk-averse seller should be willing to pay a higher risk premium to avoid the possibility of not being able to sell his shrimp for the asked price.

PREDICTION 4: *More risk-averse sellers underperform compared to less risk-averse sellers and achieve lower selling prices.*

3.4 Predicting Quality Misrepresentation

In this section, I investigate the relationship between other-regarding preferences, impatience, and quality misrepresentation. I measure quality misrepresentation by the difference between the average size of the shrimp on the top and the bottom of the pile. If the first part of prediction 1 is true, we should observe that there are differences in the degree of quality

[37] Prediction 3 is sensitive to the detection probability. With a low detection probability, it is possible that overall HC sellers do not outperform LC sellers.

misrepresentation between the different seller types: Impatient low cooperative sellers (which I denote for simplicity as I-LC, where I stands for impatient and LC for low cooperative) should misrepresent quality more than patient low cooperative sellers (P-LC) and highly cooperative sellers regardless whether they are impatient (I-HC) or patient (P-HC). If the first part of prediction 2 holds, there should be no differences in quality misrepresentation between I-HC or P-HC sellers.

I find that there are no significant differences in quality misrepresentations *overall* between LC and HC sellers or between impatient and patient sellers. LC sellers (N=18) misrepresent quality by on average 8.32 percent and HC sellers (N=15) by 6.12 percent (T-Test, t=0.46, p=0.648). I have data about the time preferences from 23 sellers that include the measure of quality misrepresentation. Impatient sellers (N=13) misrepresent quality by 8.27 percent compared to patient sellers (N=10) who misrepresent quality by 4.20 percent (T-Test, t=0.76, p=0.454). After combining other-regarding and time preferences, I find – as predicted – that there are no significant differences between I-HC and P-HC sellers (Wilcoxon Rank-Sum Test, z=-0.853, p=0.394) but significant differences between I-LC and P-LC sellers ($0.047<p<0.087$, depending on whether I use a Wilcoxon Rank-Sum Test, a T-Test or a Kolmogorov-Smirnov test of equal distributions). I-LC sellers misrepresent quality to the largest extent (11.65 percent) and significantly more than the other three seller types P-LC, I-HC and P-HC combined ($0.060<p<0.099$, depending on whether I use Wilcoxon Ranksum Test, T-Test or Kolmogorov-Smirnov test of equal distributions).[38] Note also that the difference between I-LC and P-LC is approximately three times larger than the difference between I-HC and P-HC.

Regression Table 1 presents three models with quality misrepresentation as the dependent variable. In the first model, I use a cooperativeness dummy that equals one if the seller is HC[39], an impatience dummy, and a control for whether the shrimp pile was big (it is easier to find out in small piles if the shrimp placed on top are bigger, accordingly quality misrepresentation may be more pronounced in big piles). I find that HC sellers do not misrepresent quality less than LC sellers (t=0.11, p=0.915). Impatient sellers do not significantly more misrepresent quality than patient sellers (t=0.78, p=0.443).

In the second model, I enter the interaction term impatience × cooperativeness. I observe that both the impatience dummy and the interaction term impatience × cooperativeness are

[38] Note that collecting such data is difficult. I stopped collection after I realized that some fishermen could become suspicious of this form of data collection.

[39] I use the cooperativeness dummy instead of the variable cooperativeness (which specifies the contributions from 0-10 MUs) because the variable inflation factors (which indicates multicollinearity problems) when using the variable cooperativeness are sometimes very high in the presented models (e.g. VIF of the interaction term is 9.6 with the cooperativeness variable and only 5.2 with the cooperativeness dummy) and also because the sample size is small. For more information on VIF consult e.g. Kutner et al. (2004).

marginally significant. I-LC sellers have 10.9 percent larger differences between top and bottom shrimp than P-LC sellers (t=2.08, p=0.052). There is a significant negative interaction between impatience and cooperativeness, showing that cooperativeness restrains impatient sellers from misrepresenting quality (t=-1.76, p=0.095). There are no differences between P-LC and P-HC sellers. In model 3, I further control for the beliefs in the PGE, experience, and whether the seller is a specialized shrimp seller. We observe that this model is considerably more accurate and that the impatience dummy and the interaction term are now significant at the 5 percent level. The findings in this section provide unequivocal evidence in favor of the predictions 1 and 2. Moreover, we find that specialized shrimp sellers misrepresent quality less (t=-2.15, p=0.050), which may have to do with learning since they are more experienced selling shrimp and may have figured out that placing big shrimp on top can lead to a bad reputation.

TABLE 1—DETERMINANTS OF QUALITY MISREPRESENTATION (*OLS*)

Dependent Variable	Difference in shrimp size between the top and bottom of the pile in percent		
Model	(1)	(2)	(3)
Cooperativeness	0.82 (7.55)	11.01 (6.85)	6.23 (21.16)
Impatience	3.92 (5.00)	10.93* (5.26)	10.76** (4.91)
Impatience × Cooperativeness		-16.60* (9.43)	-19.06** (8.55)
Belief in PGE			8.25 (20.70)
Experience			-0.10 (0.14)
Shrimp Seller			-10.44** (4.89)
Big Shrimp Pile	6.33 (6.69)	8.13 (5.85)	9.95 (5.99)
Constant	6.40 (3.83)	3.05 (3.82)	14.17* (7.13)
R^2	0.085	0.190	0.299
N	23	23	23

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust standard errors in parentheses. Cooperativeness, impatience, belief in PGE, shrimp seller and big pile of shrimp are all dummy variables. The cooperativeness dummy is one if the individual contributed at least four MUs in the public goods experiment. The belief in the PGE dummy is one if the individual beliefs that his group members contribute on average more than three MUs in the public goods experiment. The shrimp seller dummy is one if the seller is a specialized shrimp seller. The big shrimp pile dummy is one if the pile contains at least 50 liters of shrimp. Experience defines for how many years the seller has been catching shrimp.

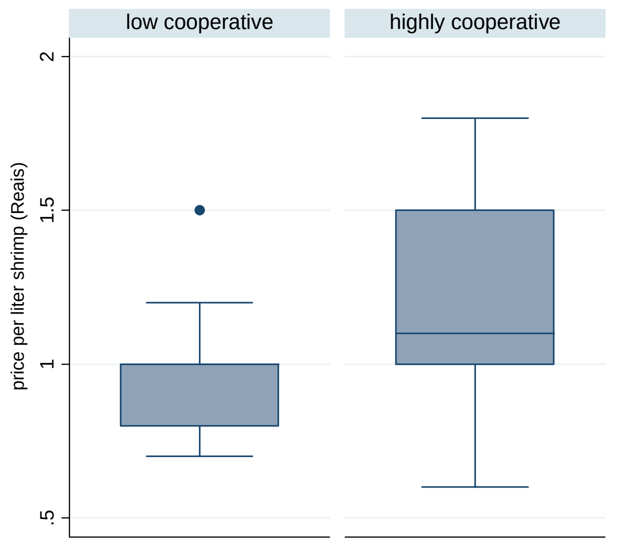
3.5 Predicting Market Performance

In this section, I relate the market performance of sellers to their other-regarding, risk, and time preferences. I measure market performance with the achieved selling price per liter shrimp. The section starts by investigating market performance *before* controlling for impatience. According to predictions 3 and 4, we should observe that highly cooperative and low risk-averse sellers achieve higher selling prices. Thereafter, I investigate market performance *after* controlling for impatience. The second part of prediction 1 says that I-LC sellers underperform compared to P-LC sellers. The second part of prediction 2 says that there are no differences between I-HC and P-HC sellers.

3.5.1 Other-regarding Preferences and Market Performance

Figure 2 illustrates the achieved selling prices per liter shrimp at the biggest market depending on the sellers' cooperativeness. We observe the achieved prices for the LC sellers who contributed little in the PGE (0-3 MUs) on the left side. The shape of the box tells us that half of the LC sellers achieve between 0.8 and 1 Reais. The median is one Real and there are only two LC sellers who get more than 1.2 Reais. On the right side, we observe the achieved prices for the HC sellers who contributed at least four MUs in the PGE. Approximately 50 percent receive between 1 and 1.5 Reais. The difference between the LC and HC sellers is highly significant (T-Test, t=-3.036, p=0.002, two-sided) and large. HC sellers achieve on average approximately 20 percent more per liter shrimp at the biggest market (1.19 vs. 0.99 Reais).

Figure 2: Other-regarding Preferences and Selling Prices



source: data from biggest market

In Regression Table 2, I investigate whether this finding holds in two OLS regressions[40] after controlling for factors that should play a role for the selling prices (e.g. risk preferences or shrimp size), and those that could do so (e.g. experience or quantity of shrimp sold). The regression model 1 only controls for risk preferences and includes market fixed effects. Regression model 2 also includes date fixed effects, shrimp size and potentially important covariates like the quantity of shrimp sold, the seller's experience, and whether the seller is specialized in selling shrimp. The outcomes in the two models with regard to cooperativeness are similar. I also find that HC sellers achieve significantly and considerably higher prices than LC sellers after controlling for shrimp size, risk preferences, and different covariates. For instance, we can see in model 2 that a HC seller receives 0.136 Reais more per liter shrimp compared to a LC seller ($t=2.11$, $p=0.038$).[41] Therefore, I find support for prediction 3.

As expected, shrimp size is positively related to the selling price ($p<0.001$); the seller's experience and specialization in selling shrimp also matter ($p<0.040$). The quantity of shrimp sold does not significantly predict the price per liter shrimp ($p=0.556$), suggesting that there are no important quantity reductions at our shrimp markets. Note also (not shown) that cognitive skills do not significantly predict market performance. For instance, if I introduce cognitive skills as an additional independent variable in model 2, it is uninformative (Observations: 68, coefficient=0.021, $p=0.508$ in model 1). Sellers with better cognitive skills measured by the calculation task do not perform better.

[40] Since the dependent variable price per liter shrimp is not normally distributed, using a generalized linear model (GLM) might be more appropriate. I also used the GLM and find that results do not change much. If at all, the findings with regard to cooperativeness and risk-aversion get more pronounced and more robust.

[41] If I use the cooperativeness variable which specifies how many of the ten MUs were contributed in the regression instead of the cooperativeness dummy, I get similar results. For instance, every MU contributed more is associated with a 0.018 Reais increase in selling prices in model 2 ($t=2.13$, $p=0.037$).

TABLE 2—DETERMINANTS OF MARKET PERFORMANCE I (*OLS)*

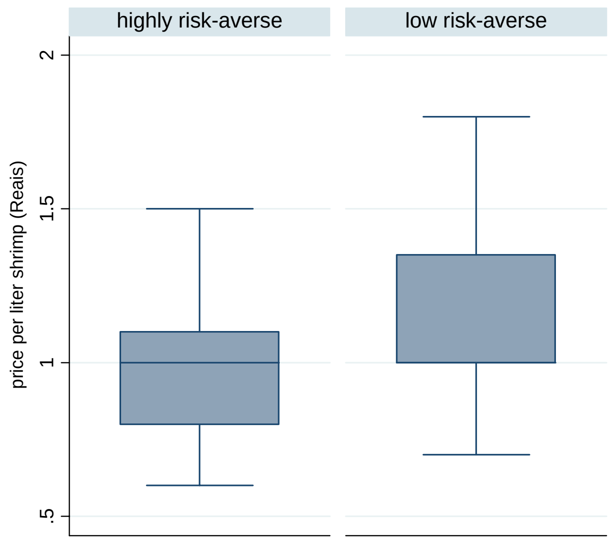
Dependent Variable	Selling price per liter shrimp	
Model	(1)	(2)
Cooperativeness	0.092**	0.136**
	(0.046)	(0.065)
Risk-Aversion	-0.130***	-0.154***
	(0.047)	(0.046)
Shrimp Size (Quality)		0.153***
		(0.044)
Shrimp Color		-0.088
		(0.053)
Belief in PGE		-0.087
		(0.054)
Experience		0.003**
		(0.001)
Shrimp Seller		0.098**
		(0.047)
Liter Shrimp sold		-0.000
		(0.000)
Market Fixed Effects	yes	yes
Date Fixed Effects	no	yes
Constant	0.961***	0.447***
	(0.038)	(0.168)
R^2	0.814	0.855
N	99	93

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust standard errors in parentheses. High cooperativeness, high risk-aversion, color of shrimp, belief in PGE and shrimp seller are all dummy variables. The cooperativeness dummy is one if the individual contributed at least four MUs in the public goods experiment. The risk-aversion dummy is one if the individual invested at least three MUs in the RAE. Color of shrimp dummy is one if the shrimp are colored. The belief in the PGE dummy is one if the individual beliefs that his group members contribute on average more than three MUs in the public goods experiment. The shrimp seller dummy is one if the seller is a specialized shrimp seller. Experience defines for how many years the seller has been catching shrimp.

3.5.2 Risk-aversion and Market Performance

The next boxplots in Figure 3 illustrate the selling prices at the biggest market depending on the sellers' level of risk-aversion. We observe the achieved selling prices for the highly risk-averse fishermen who risked a maximum of two MUs in the RAE on the left side. The shape of the box shows that approximately half of the highly risk-averse fishermen achieve between 0.8 and 1.1 Reais and the mean is 0.99 Reais. In contrast, the less risk-averse fishermen depicted on the right side who risked at least three MUs in the RAE achieve significantly higher prices (T-Test, t=-2.19, p=0.033). Approximately 50 percent of the low risk-averse sellers receive between 1 and 1.4 Reais and the mean is 1.13 Reais.

Figure 3: Risk-aversion and Selling Prices



Source: data from biggest market

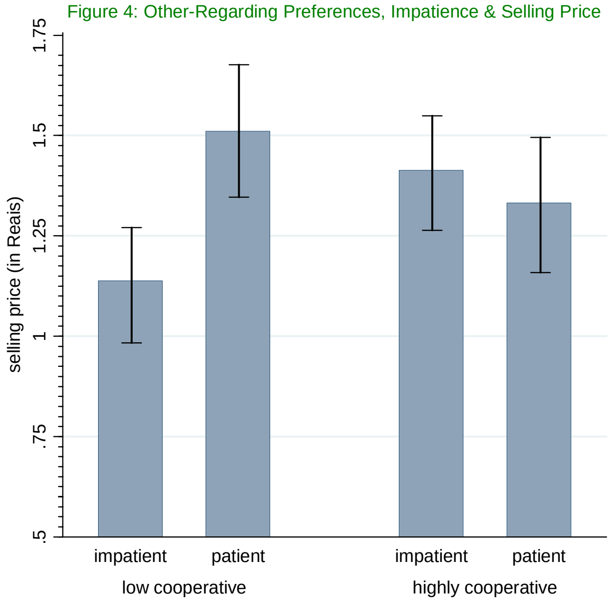
It becomes further clear in Regression Tables 2 (and 3) that the level of risk-aversion significantly hampers market performance. The level of risk-aversion is highly predictive, even after controlling for other covariates like cooperativeness, shrimp size, or impatience. In model 2 of Table 2, for instance, we observe that highly risk-averse sellers achieve 0.154 Reais less per liter shrimp than low risk-averse sellers (t=3.33, p=0.001), indicating a substantial impact of risk-aversion on market performance.[42] Thus, I provide evidence in favor of prediction 4.

[42] I attain similar results if I use the risk-aversion variable which specifies how many of the ten MUs were not invested in the lottery instead of the highly risk-averse dummy in the regression. In model 2, for instance, each fewer invested MU is associated with a 0.028 Reais decrease in selling prices (t=2.59, p=0.012).

3.5.3 *Other-regarding Preferences, Impatience, and Market Performance*

Finally, I investigate the effect of other-regarding preferences *after* controlling for impatience. I have a total of 51 observations which include the selling price and the laboratory measures of other-regarding and time preferences. Figure 4 illustrates the achieved selling prices for four types: I-LC, P-LC, I-HC, and P-HC. The two bars on the left indicate the average selling prices for the LC sellers. While there is no significant difference between impatient (mean = 1.30 Reais) and patient sellers (1.41 Reais; T-test, t=-0.738, p=0.232, one-sided) overall, we can clearly observe a big difference between I-LC and P-LC sellers. I-LC sellers receive on average 1.14 Reais whereas P-LC sellers receive approximately 30 percent more (1.51 vs. 1.14 Reais, t=-1.67, p=0.055, one-sided; Wilcoxon Rank-Sum Test p=0.079). There is no such negative effect of impatience for the HC sellers. I-HC and P-HC sellers receive approximately the same (1.33 vs. 1.41, t=0.39, p=0.649, one-sided).

Figure 4: Other-Regarding Preferences, Impatience & Selling Price



note: standard error bands shown

Next, I present in Regression Table 3 two models for observing whether we find the same pattern after controlling for size and the other covariates I used in Regression Table 2. As suggested in Figure 4, we observe an interesting relationship between impatience and cooperativeness for market performance. First, the cooperativeness dummy, which now shows the effect of high cooperativeness for patient sellers, is insignificant and negative, suggesting that high cooperativeness does not improve performance for patient sellers. Second, the impatience dummy, which presents the effect of impatience for the LC sellers, is highly significant. An I-LC seller receives between 0.229 and 0.264 Reais less per liter shrimp compared to a P-LC seller (significant at 1 percent level), depending on the model specification. Third, the interaction between impatience and cooperativeness is positive and highly significant, showing that high cooperativeness helps impatient sellers to improve their performance ($p<0.006$).[43] Size and risk-aversion are the only other variables that predict market performance significantly ($p<0.004$). Thus, I find support for predictions 1 and 2, and that the overall effect that HC sellers achieve higher selling prices than LC sellers is mediated by I-LC sellers.

[43] I obtain similar results if I use the variable cooperativeness and not the cooperativeness dummy. The interaction impatience x cooperativeness is highly significant in both models ($p<0.011$).

TABLE 3—DETERMINANTS OF MARKET PERFORMANCE II
(*OLS)*

Dependent Variable	Selling price per liter shrimp	
Model	(1)	(2)
Cooperativeness	-0.101 (0.061)	-0.133 (0.090)
Impatience	-0.229*** (0.067)	-0.264*** (0.070)
Impatience × Cooperativeness	0.271*** (0.093)	0.368*** (0.098)
Risk-Aversion	-0.191*** (0.060)	-0.196*** (0.064)
Shrimp Size (Quality)		0.222*** (0.069)
Shrimp Color		0.009 (0.074)
Belief in PGE		-0.055 (0.073)
Experience		0.003 (0.003)
Shrimp Seller		0.003 (0.066)
Liter Shrimp sold		-0.000 (0.000)
Market Fixed Effects	yes	yes
Date Fixed Effects	no	yes
Constant	1.140*** (0.058)	0.629** (0.243)
R^2	0.899	0.924
N	51	48

Notes: *** 99-percent significance, ** 95-percent significance; * 90-percent significance. Robust standard errors in parentheses. High cooperativeness, high risk-aversion, impatience, color of shrimp, belief in PGE and shrimp seller are all dummy variables. The cooperativeness dummy is one if the individual contributed at least four MUs in the public goods experiment. The risk-aversion dummy is one if the individual invested at least three MUs in the RAE. Color of shrimp dummy is one if the shrimp are colored. The belief in the PGE dummy is one if the individual beliefs that his group members contribute on average more than three MUs in the public goods experiment. The shrimp seller dummy is one if the seller is a specialized shrimp seller. Experience defines for how many years the seller has been catching shrimp.

3.6 Does Market Exposure affect Individual Preferences?

I showed in the last section that certain (combinations) of individual preferences are associated with better market performance. A potential follow-up question is whether in our setting individual preferences affect (performance in) markets or whether (performance in) markets affect individual preferences. My data allows investigating whether sellers who are less exposed to markets, i.e. sellers who do not always sell their shrimp at markets, are endowed with different individual preferences. If the studied markets affect individual preferences, according to my previous findings, there are reasons to believe that sellers who are more exposed to markets are more cooperative, less risk-averse, and less impatient because these are the individual preferences associated with a better market performance.[44] [45] I find that this is not the case. Sellers who always sell their catch at markets are, if at all, less cooperative and contribute on average 3.51 MUs in the PGE compared to the 3.96 MUs the sellers who do not always sell their catch on markets contribute ($t=1.12$, $p=0.261$). Sellers who always sell their shrimp at the markets are also not less impatient; if anything, they are rather more impatient (Spearman Rank Correlation, $r=-0.161$, $p=0.112$).[46]

Market exposure also does not affect the level of risk-aversion. Fishermen who always sell their catch at markets risk 2.99 MUs on average in the RAE, approximately the same amount that fishermen risk who do not always sell their catch on markets (3.06 MUs, $t=0.23$, $p=0.819$). Therefore, the data suggests that the exposure to shrimp markets does not affect individual preferences because sellers who are not or less exposed to the markets have no significantly different sets of individual preferences.

[44] Of course, having these preferences may be also beneficial when selling shrimp not directly at the shrimp market. However, since the competitive pressure seems to be high at the shrimp markets, it is plausible that the shrimp markets affect individual preferences in the suggested direction.

[45] It could also be that there is a selection which type of fisherman is more likely to sell his shrimp at markets. If selection plays a role, we should observe that sellers who are more exposed to the markets are more cooperative, less risk-averse, and more patient. Since this is not the case, selection does not appear to play an important role.

[46] One may argue that more market exposure reduces the probability that there are I-LC sellers at the markets (see sections 3.3 to 3.5). This is not the case, however. As many as 32.35 percent (N=22) of the sellers who always sell their catch at markets are I-LC, which is more than the fraction of P-LC (17.64 percent) and the fraction of P-HC sellers (17.64 percent). The remaining 32.35 percent are I-HC sellers.

3.7 Conclusion

The goal of this paper is to examine how other-regarding, risk, and time preferences are related to behavior and performance in naturally occurring markets where reputation and asymmetric information play an important role. I present and test a simple model based on the idea that other-regarding preferences can restrain impatient sellers from yielding to their temptation to engage in uncooperative behavior and thus prevent them from hampering their market performance. The model is supported by the data. Impatient less cooperative sellers misrepresent quality more compared to the other seller types. Moreover, I find that less cooperative sellers receive considerably lower prices for products of similar quality compared to more cooperative sellers, and that time preferences mediate the difference in performance. Less cooperative sellers only underperform compared to the other seller types if they are impatient. Concerning risk-aversion, I find new evidence from the field that risk-aversion is negatively related to market performance.

These findings provide insights into the role of other-regarding preferences and corroborate the relevance of social preference and reciprocity theories in understanding behavior and outcomes in naturally occurring markets. The studied markets share features with many other typical market settings; nevertheless, it would be interesting to investigate the relationship between individual preferences and market performance in markets with different characteristics. For instance, how are individual preferences related to individual performance on spot markets? According to the presented model, other-regarding preferences play a less important role in markets where reputation building is not crucial.

The study shows the importance of investigating combinations of different individual preferences to better understand economic outcomes in natural markets. It would also be interesting to observe whether there are other preference combinations that further explain market performance. It could be, for instance, that risk-loving selfish sellers also significantly underperform compared to risk-loving other-regarding sellers. In this paper, I have shown that other-regarding preferences can restrain impatient individuals from uncooperative behavior. However, studying whether impatience deters individuals with insufficiently pronounced other-regarding preferences from being cooperative also seems worthwhile. In any case, developing a more general behavioral model that accounts for combinations of different preferences within one individual would be very interesting.

3.A Appendix Tables and Figures

TABLE a—SUMMARY STATISTICS
(Means)

	Mean	Observations
Market Data		
Selling price (in Reais)	1.272 (0.051)	99
Shrimp size (in cm)	2.966 (0.042)	99
Liter shrimp sold	83.516 (8.2478)	93
Big shrimp pile	0.441 (0.052)	93
Shrimp color	0.444 (0.050)	99
Difference between top and bottom shrimp (in %)	7.320 (2.346)	33
Difference between shrimp size guess and actual shrimp size (in cm)	0.335 (0.147)	37
Laboratory Data		
Contributions in PGE	3.667 (0.189)	216
Cooperativeness	0.486 (0.034)	216
Belief in PGE	0.551 (0.034)	216
Impatience	0.586 (0.050)	99
Risked MUs in RAE	3.018 (0.150)	216
Risk-aversion	0.398 (0.033)	216
Survey Data		
Experience (in years)	17.723 (0.806)	215
Specialized shrimp seller	0.750 (0.029)	216
Cognitive skills	1.873 (0.060)	150

Notes: Standard errors in parentheses. Sample sizes differ because data come from different sources. Market data was collected at the market in April-May 2008. Big shrimp pile is a dummy that equals one if the pile contains at least 50 liters. Shrimp color is dummy that equals one if the shrimp were colored. Laboratory data was collected in village meetings in April 2008 and 2006. Cooperativeness is a dummy that equals one if at least four MUs were contributed. Belief in PGE is a dummy that equals one if the participant believed the other players contribute on average at least 3.5 MUs. Impatience is a dummy that equals one if the participant preferred one bottle mineral water immediately over two bottles the next day. Risk-aversion is a dummy that equals one if the participant invested less than three MUs in the RAE. The survey data was collected in April and August 2008. Cognition skills define how many of the three calculation questions were answered correctly. Specialized shrimp seller is a dummy that equals one if the seller is a specialized shrimp seller.

Figure a: Picture of Shrimp Market and Shrimp Piles

Figure b: Shrimp Sizes at T and T+1

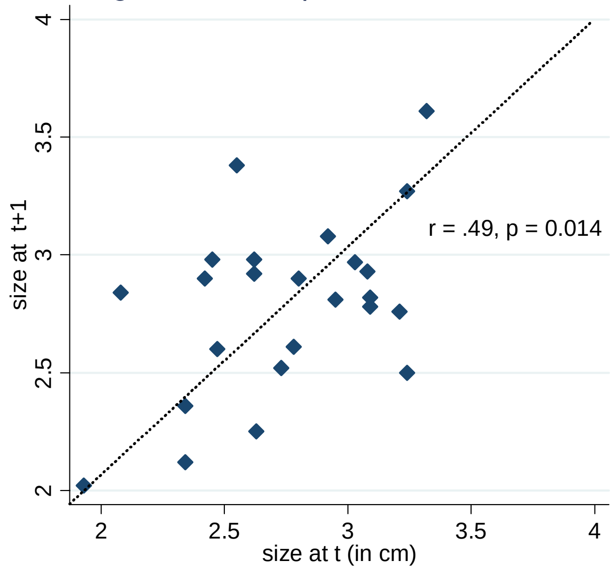


Figure c: Achieved Selling Prices

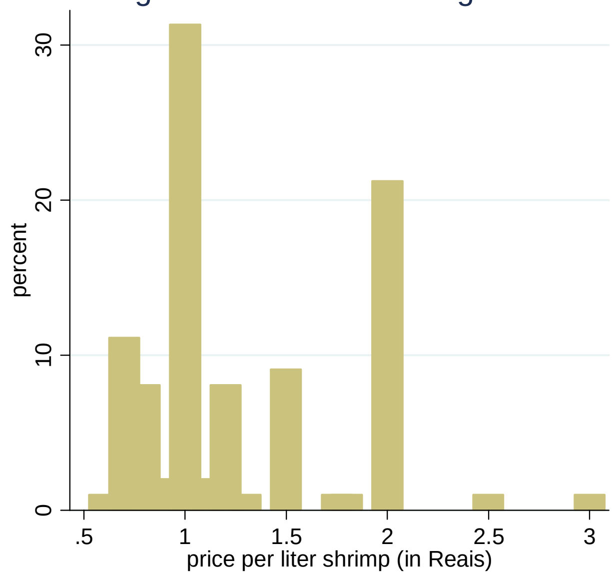


Figure d: Average Shrimp Size

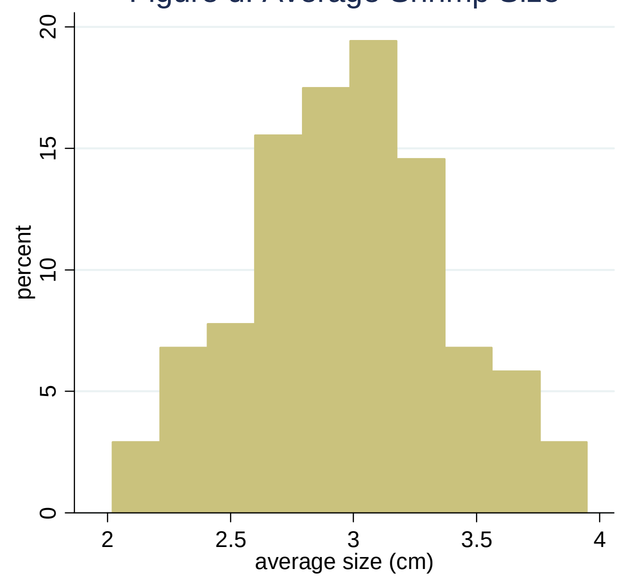
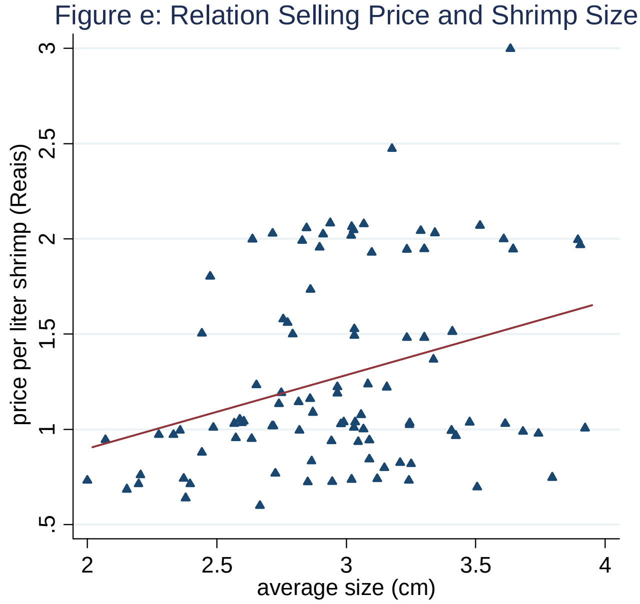


Figure e: Relation Selling Price and Shrimp Size



Note: line illustrates linear prediction. Triangles are slightly jittered for illustrative reasons.

Chapter 4

AN EXPLORATION OF THIRD AND SECOND PARTY PUNISHMENT IN TEN SIMPLE GAMES

Chapter Overview

This paper identifies the motives behind punishment from unaffected third parties and affected second parties using a within-subject design in ten simple games. We apply a classification analysis and find that a parsimonious model assuming that subjects are either inequity averse or selfish best explains the pattern of punishment from both third and second parties. Despite their unaffected position, we do not find that third parties punish in a more impartial or normative manner.

4.1. Introduction

Third parties play a crucial role in many institutions: They serve in courts, as referees or arbitrators.[47] The US legal system, for instance, relies on their judgment in juries when it comes to the application of sanctions. Third parties are also important with regard to informal sanctions (Homans, 1961) and, in fact, their interventions seem to be essential in the explanation of norm enforcement, as they are often more numerous than affected second parties (Bendor and Swistak, 2001) or the only parties present (Greif, 1993, 1994), and hence their sanctions are potentially more damaging than those from second parties.

Despite their importance, little is known about how third parties sanction others. In particular, it is unclear whether third parties sanction in a different manner than second parties. In principle, third parties might sanction in a more impartial, "normative", and controlled manner, and less egocentrically (Fehr and Fischbacher, 2004). Adam Smith apparently had this idea in mind when he introduced the concept of the "impartial spectator" in his *Theory of Moral Sentiments*, a party who is not personally affected, making decisions from beyond the limitations of egocentric biases. In fact, the prevalence of institutions that rely on third parties implies that they are likely to make more appropriate decisions. However, it also seems plausible that even third parties cannot completely eliminate egocentric biases (Ross et al., 1976; Babcock et al., 1995). The concerns about the selection of jury members in many law cases suggest that third parties can make very inappropriate decisions in the context of sanctioning (e.g. Kennedy, 1997).

Recent models of other-regarding preferences propose competing explanations for third and/or second party punishment. Theories based on inequity-aversion (Fehr and Schmidt 1999, Falk and Fischbacher 2006) predict punishment of richer co-players if that reduces the payoff distance, while reciprocity theories predict punishment of an individual if she harmed the aggressor (Rabin 1993, Dufwenberg and Kirchsteiger 2004; Cox et al., 2007). Further, Bolton and Ockenfels (2000) predict punishment of any co-player if that brings the aggressor's relative payoff closer to the average relative payoff, Levine (1998) posits the existence of spiteful types who punish indiscriminately and type-reciprocal agents who punish selfish or spiteful co-players, and López-Pérez (2008) predicts punishment of norm deviators.

This paper applies a within-subjects experimental analysis and the classification method by El-Gamal and Grether (1995), with two key objectives: (i) to study and compare the motives

[47] We say that a player C is a third party with respect to a player A if her material payoff does *not* depend on the decisions of A (note however that it is possible that the material payoff of A depends on the decisions of C; for instance, it could be the case that C sanctions A and thus reduces her material payoff). We also say that a player B is a second party with respect to A if her material payoff depends on A's decisions.

behind third and second party punishment, and (ii) to provide a stringent test of recent models and determine which one best accounts for punishment in a large range of situations.[48] Our paper suggests several insights into the key motives behind punishment. *First*, we find that both third and second party punishment is predominantly targeted towards richer co-players. *Second*, the classification analysis shows that a model assuming that subjects are either inequity-averse or selfish captures the *occurrence* of third *and* second party punishment across our ten games better than any other equally parsimonious alternative. While models that also include small fractions of spiteful (for third and second parties) and reciprocal (for second parties) types are slightly more accurate, they come at the cost of increased complexity. *Third*, we observe that the *strength* of punishment depends heavily on the size of the payoff disadvantage (in 3P and 2P). *Fourth*, and contrary to the idea that third parties are less "infected" by egocentric or "non-normative" motives than second parties, we observe that third party punishment generally resembles second party punishment and find no support for the conjecture that third parties are impartial and enforce (informal) rules in a normative manner.

A large body of experimental research shows that subjects are often willing to spend money to reduce another player's payoff – i.e. to punish her – even if no future benefits can follow from this behavior. In the ultimatum game, responders frequently punish proposers for making unfair offers (Güth et al. 1982, Camerer and Thaler 1995, Roth 1995), while non-contributors are often punished in public goods games with a punishment stage (Fehr and Gächter 2000). However, this literature has a completely different focus than our study because it is restricted to second party punishment and because the analyzed games are not well suited for discriminating the motives behind punishment. In the ultimatum game, responders might reject offers due to inequity-aversion, reciprocity, spite, or to punish a violation of an equity norm, while punishment in the public goods game can be explained in terms of inequity-aversion, reciprocity, spite, or as a reaction to a transgression of a cooperation norm.

The studies by Falk et al. (2005) and Dawes et al. (2007) have provided progress in our understanding of *second-party* punishment, although considerable uncertainties remain regarding the main motives. First of all, the two studies come to different conclusions. Falk et al. (2005) find that "retaliation seems to be the most important motive behind fairness-driven informal sanctions" (*ibid*, p. 2017) whereas Dawes et al. (2007) emphasize the importance of the "egalitarian motives". Moreover, these studies do not allow for perfect discrimination between some motives. For instance, the results from Falk et al. (2005) are not only consistent with reciprocity (i.e., retaliation) but also with a model predicting punishment of players who

[48] The reader interested in similar approaches to understand behavioral decision rules may consult Engle-Warnick (2003).

deviate from a norm of cooperation/efficiency (see López-Pérez, 2008). In turn, reciprocity might explain part of the results in Dawes et al. (2007). In this study, subjects were placed in small groups and allocated a randomly determined sum of money which could be used to reduce the income of their co-players'. Since all players could engage in damaging, they could damage conditional on their expectations about the damaging from the other subjects, i.e. retaliate (see Zizzo, 2003).

In addition, the studies by Falk et al. (2005) and Dawes et al. (2007) pay little attention to players' heterogeneity because they consider very few games and use a between-subject design. This seems to be a limitation because punishment is very likely caused by multiple motivational forces. On the one hand, different players might have different reasons to punish in the same game –the evidence from Falk et al. (2005) is indeed consistent with this. On the other hand, the same player might punish for one reason (say, inequity-aversion) in one game, and for another reason (say, reciprocity) in a different game. Therefore, we believe that it is crucial to investigate which forces are *relatively* more powerful for each player across a large range of games and to provide an analysis that allows classifying subjects as (predominantly) inequity-averse, reciprocal, etc.

Few studies address *third party* punishment (e.g. Zizzo 2003; Carpenter and Matthews, 2005; Charness et al., 2008) and Fehr and Fischbacher (2004) is the only study which compares it to second party punishment. The authors report that third parties punish unfair allocation choices in a dictator game and defectors in a prisoner's dilemma game, although less strongly than second parties do. However, it remains unclear why third parties punish in these games (it could be because they punish violations from norms of cooperation/equity, but also because of inequity-aversion, spite, or because they are type-reciprocal á la Levine, 1998) and why they punish less than second parties (this might be an artifact of their experimental design, as the payoff disadvantage was larger between first and second parties than between first and third parties). To the best of our knowledge, our study is the first to (i) identify the motives behind third party punishment in a comprehensive design and (ii) investigate whether third parties are impartial and less prone to punish in an egocentric or "non-normative" manner than second parties.

Our results indicate that it can be misleading to assume that third parties are impartial, make less egocentric choices, and enforce (informal) rules in a normative manner. In addition, this study gives important implications for the further development of recent theories of other-regarding preferences. Inequity-aversion appears to be an indispensable factor in explaining the occurrence and strength of third and second party punishment, while reciprocity and spite play an important although *relatively* minor role in explaining the occurrence of second party punishment. In particular, we believe that reciprocity should not be used alone in predicting punishment in general, as it fails to explain any third party punishment.

The rest of the paper proceeds as follows. The next section presents the experimental design and procedure. Section 4.3 describes our research hypotheses. In section 4.4, we report the results from the classification analysis and study which factors affect the occurrence and strength of second and third party punishment. The section 4.5 concludes.

4.2 Experimental Design and Procedures

There are two treatments in our experimental design: A second party punishment treatment (2P) and a third party punishment treatment (3P). Participants in 2P play ten two-player games, while participants in 3P play ten three-player games. All these games have a two-stage structure. In the first stage of both treatments, one player (the *first party*) chooses between a left-hand and a right-hand allocation of payoffs between herself and another player (the *second party*). Table 1 shows the two allocations available in each game (we explain in section 4.3 why we chose these particular allocations). They are identical in 2P and 3P and presented in points (10 points = 1 Swiss Franc).

TABLE 1—THE ALLOCATIONS IN THE 10 GAMES

		Game									
		1	2	3	4	5	6	7	8	9	10
Allocation	Left	(150,150)	(100,100)	(560,60)	(150,90)	(220,260)	(280,240)	(250,80)	(100,100)	(250,150)	(250,150)
	Right	(590,60)	(50,530)	(120,140)	(50,630)	(220,400)	(390,240)	(80,250)	(50,150)	(110,290)	(330,70)

The second stage differs in the two treatments. In any game of 2P, the second party can spend points out of her allocation share to reduce the first party's payoff –i.e., to punish her. In any game of 3P, a third player (the *third party*) can punish the first *or/and* the second parties, while the second party in 3P makes no decision, i.e. she is a "bystander". The third party is endowed with 200 points in each allocation of each game meaning the first party's choice never affects her payoff in the first stage. The punishment technology is the same in 2P and 3P: Up to 50 points can be used to punish and each point spent reduces the payoff of the punished player by three points. Hence, if the first party chooses the allocation $(x_{FP},\ x_{SP})$ in a game in 2P and the second party punishes her with $0 \leq p \leq 50$ points, the first party's payoff in that game is $x_{FP} - 3p$ and the second party's payoff is $x_{SP} - p$. In 3P, if the first party chooses allocation

(x_{FP}, x_{SP}) in a game and the third party punishes her with p_1 points and the second party (the bystander) with p_2 points ($p_1 + p_2 \leq 50$), the payoffs in this game are $x_{FP} - 3p_1$ for the first party, $x_{SP} - 3p_2$ for the second party, and $200 - p_1 - p_2$ for the third party.

We ran eight sessions where we observed a total of 3100 punishment decisions. Each session proceeded as follows. Subjects were randomly assigned to be a first or second party (or third party in 3P) and anonymously matched in groups of two (in 2P) or three (in 3P). Each subject received instruction sheets (dependent on role and treatment) which explained the extensive form of the games (without giving information about the payoff constellations of the ten games). Subjects had to fill out control questions to make sure that they understood the rules. We used neutral language and avoided terms such as “punishment”. Every subject always played the ten games in the same role and no subject participated in both treatments. The ten games were presented one at a time, and the order in which they were played was randomly predefined for each group. At the end of each game, we asked first parties in 2P and 3P (and second parties in 3P) about their expectations of punishment in each allocation. Subjects were never told about their counterparts’ previous choices to prevent repeated game effects. After the subjects played the ten games, only one game was randomly selected for payment in order to prevent income effects.[49]

In the eight sessions, we employed the strategy method to elicit the punishment behavior in the second stage, i.e. the subjects had to indicate for both allocations in each of the ten games the number of points (0–50) they wanted to assign to the other subject(s). In principle, the strategy method might induce a different behavior than the specific response method, where subjects face given, known choices for one allocation or the other.[50] However, Falk et al. (2005) investigate this issue and find no differences in subjects’ punishment patterns, although the strength of punishment is somewhat lower overall with the strategy method. Thus, the existing evidence suggests that the strategy method does not affect the pattern of punishment, but might possibly lead to an under-representation of actual punishment.

The key reason for using the strategy method was to prevent subjects from receiving any feedback about the first party's choices in any of the ten games, something that would lead to serious confounds: Punishers’ mood could change depending on the first party’s prior behavior, and this could generate order or history effects which would severely complicate the data

[49] It could be argued that this dilutes monetary incentives because subjects make more decisions for the same amount of money. However, a meta-study by Camerer and Hogarth (1999) suggests that this is not the case.

[50] For other decisions than punishment, there is evidence of no systematic differences in behaviour between the strategy and specific response method (Cason and Mui, 1998; Brandts and Charness, 2000; Falk and Kosfeld, 2006).

analysis.[51] In our view, the use of the strategy method seems unavoidable for the study of punishment behavior with a within-subjects design and a large set of games (unless the researcher has access to huge samples in order to control for order effects). Additionally, it maximizes the amount of statistical data gathered.

The experiment was conducted with the Z-tree software (Fischbacher, 2007) and the participants were recruited with the software "ORSEE" (Greiner, 2004). 255 subjects participated in our experiment, 90 in 2P and 165 in 3P, that is, we observed 45 second and 55 third parties. Most subjects were students from different disciplines of the University of Zurich or the Swiss Federal Institute of Technology in Zurich (9 percent of them came from the faculty of economics and management). They earned on average 30 Swiss Francs (around $ 24) which included a show-up fee of 10 Swiss Francs (this fee could be accordingly reduced if one subject got a negative point score as a result of heavy punishment, although this never happened). The sessions lasted approximately 60 minutes. The instructions for both treatments are in the appendix in section A.2.1 and A.2.2.

4.3 Research Hypotheses

Our main research objective is determining the key forces that explain the occurrence (and strength) of third and second punishment. Note in this regard that, as subjects in our experiment get no feedback and are paid for their choices in just one game, we can analyze each single game as a one-shot situation. Therefore, our selection of games enables us to rigorously test numerous motives for the occurrence of costly punishment.[52] The standard model assuming self-interested players predicts that neither third nor second parties should ever punish in any of our games. In contrast, existing models of other-regarding preferences predict punishment in our games under certain circumstances (we later provide a more formal description that includes other alternative models):

---

[51] As an illustration, consider a second party who first plays against an "unkind" first party and gets angry as a result. This negative emotional state could affect her posterior behavior, even if the new opponent (players should be re-matched when using the specific response method in order to prevent repeated game effects) makes a "kind" choice. In this regard, Fehr and Fischbacher (2004) report spillover effects when using the specific response method in their two treatments where participants played two games with re-matching. To keep this spillover from contaminating their results, they had to restrict the analysis to the games that were played first.

[52] Our focus in this paper is on the causes of punishment. In a companion paper, we provide a detailed analysis of several consequences of inequity-averse punishment, like (i) the punishment of socially efficient (or even Pareto efficient) choices, (ii) the sanctioning of by-standers, and (iii) the absence of punishment when absolute payoff equality is achieved.

- Inequity-aversion (Fehr and Schmidt, 1999; Falk and Fischbacher, 2006): A second party (in 2P) or a third party (in 3P) punishes in any allocation of the ten games if she gets a smaller payoff than another co-player and she is sufficiently averse to disadvantageous inequity.[53]
- ERC (Bolton and Ockenfels, 2000): This model makes the same predictions on the occurrence of punishment in 2P as inequity-aversion. In 3P, it predicts punishment of the first *and/or* the second party at any allocation if the third party can thus bring her relative payoff closer to 1/3, the equitable relative payoff in three-player games.
- Reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Cox et al., 2007): A second party in 2P punishes if she is sufficiently reciprocal and the first party chooses the allocation of the game giving the second party the strictly lowest payoff –i.e. if the first party harmed the second party. These models predict no punishment in 3P, as the third party's endowment is always 200.
- Spite (inspired by Kirchsteiger, 1994 and Levine, 1998): Spiteful second and third parties should punish a co-player at all allocations.[54]
- Anti-greed (inspired by Levine, 1998): A second or third party of this type punishes the first party if the latter chose the allocation maximizing her own money payoff, the intuition being that second and third parties punish selfish or greedy first parties.
- Competitiveness (inspired by Levine, 1998): This is the opposite of inequity-aversion, that is, second and third parties punish in order to increase an already positive income difference.[55]
- Efficiency (inspired by López-Pérez, 2008): A second or third party of this type punishes the first party if the latter chose the least efficient allocation of the game (i.e., the allocation with the smallest sum of payoffs); the intuition being that second and third parties punish deviations from a norm of social efficiency.
- Equity (inspired by Elster, 1989 and López-Pérez, 2008): A second or third party of this type punishes the first party if the latter chose the least equitable allocation of the game (i.e., the allocation with the largest distance between players' payoffs); the intuition being that second and third parties punish deviations from a norm of equity.

[53] Of course, all the models that we cite in this section predict punishment under certain parameterizations. If the players differ in their parameters, therefore, only some of them might punish in our games. However, we do not need to be precise a priori on the distribution of parameters in the population because indeed this is part of what we aim to clarify with the classification analysis.

[54] This might seem a very stringent prediction, but recall that our experimental design was such that only one allocation was chosen for payment in both treatments. A spiteful type would, therefore, punish in all allocations.

[55] Although Levine allows for the existence of spiteful types that should punish indiscriminately, and estimates that around 20 percent of the population correspond to this type, he suggests later that "one explanation of spite is that it is really "competitiveness," that is, the desire to outdo opponents" (Levine 1998, p. 614).

Table 2 presents the theoretical predictions of the models in each of our ten games for both third and second party punishment. In the first two columns, we can see the two allocations in each of the ten games. The following two columns indicate for the 2P treatment which theories predict punishment in the left-hand and right-hand allocation of each game. As an illustration of these predictions, take game 6 (280/240 vs. 390/240). An inequity-averse second party punishes the first party either if he chooses the left-hand or the right-hand allocation, because she always gets a lower payoff than him, while a reciprocal individual punishes neither allocation because she cannot be harmed in this game. In contrast, an individual who punishes deviations from a norm of equity would punish the choice of the right-hand allocation (390/240) because it is more inegalitarian than the alternative (280/240), and an individual who punishes deviations from a norm of social efficiency would punish the left-hand allocation (280/240) because the joint payoff is bigger in the alternative allocation (390/240). The next four columns illustrate the punishment predictions for 3P. We first show the punishment predictions for the first party and then the punishment for the bystander.[56]

In addition, the last four columns of Table 2 categorize the games according to four criteria: (1) JPM, i.e. whether a joint-payoff maximizing allocation exists in the respective game, (2) PARETO, i.e. whether a Pareto-dominant allocation exists in the respective game, (3) STRICT, i.e. whether a strictly equal allocation exists in the respective game and, (4) EQUITY, which specifies whether the party who can punish is monetarily disadvantaged by the allocation that is less equal. We believe that these four criteria can be important to understand punishment and help to omit potential biases towards certain theories. It is possible, for instance, that second and/or third parties are less willing to punish the choice of unequal allocations if they are socially or Pareto efficient: Maybe some third and/or second parties are willing to accept a small disadvantage for a great advantage of the other player. This seems especially plausible as several studies suggest that some deciders choose socially efficient allocations even at their own material disadvantage (Charness and Rabin 2002, Engelmann and Strobel 2004; Fehr et al. 2006). In other words, inequity-aversion might be less influential at any social or Pareto efficient allocation. That STRICT is important has been suggested in abundant investigations in negotiation and mediation that recommend reaching strictly equal outcomes because the involved parties often accept them and do not want to deviate from them (e.g. Thompson 2005). In this respect, inequity-aversion might be more influential if the game has a strictly equal allocation –in line with this, Güth et al. (2001) find that in ultimatum games an unfair

[56] If a theory predicts that the third party is indifferent between punishing the first party or the by-stander at one allocation, we take as compatible with the theory the punishment of any of those two parties.

offer is more often rejected if the alternative is a strictly equal instead of a slightly unequal offer. The category EQUITY complements the previous category.

In summary, our comprehensive design allows us to discriminate between many different models and takes care of potential confounds. Moreover, the comparatively large number of games makes inferences possible with respect to how consistently individuals follow each behavioral rule, thus providing a stress test of models of other-regarding preferences.

TABLE 2— THEORETICAL PUNISHMENT PREDICTIONS AND GAME CHARACTERISTICS

	Game			Theories predicting punishment in 2P		Theories predicting punishment for the first party in 3P		Theories predicting punishment for the by-stander in 3P		Game Characteristics			
	Allocation			Allocation		Allocation		Allocation		JPM	Pareto	Strict	Equity
	Left		Right	Left	Right	Left	Right	Left	Right				
1	(150,150)	vs.	(590,60)	EF, C	IA, R, EQ, AG	EF, C	IA, EQ, AG	C	C	yes	no	y	y
2	(100,100)	vs.	(50,530)	R, EF, AG, C	EQ, C	EF, AG, C	ERC, EQ, C	C	IA, ERC	y	n	y	n
3	(560,60)	vs.	(120,140)	IA, R, EQ, AG	EF, C	IA, EQ, AG	EF, C	C	C	y	n	n	y
4	(150,90)	vs.	(50,630)	IA, R, EF, AG	EQ, C	EF, AG, C	EQ, C	C	IA	y	n	n	n
5	(220,260)	vs.	(220,400)	R, EF, C	EQ, C	IA, ERC, EF	IA, EQ	IA, ERC	IA	y	y	n	n
6	(280,240)	vs.	(390,240)	IA, EF	IA, EQ, AG	IA, ERC, EF	IA, EQ, AG	IA, ERC	IA	y	y	n	y
7	(250,80)	vs.	(80,250)	IA, R, AG	C	IA, AG	C	C	IA	n	n	n	-
8	(100,100)	vs.	(50,150)	R, AG, C	EQ, C	AG, C	EQ, C	C	C	n	n	y	n
9	(250,150)	vs.	(110,290)	IA, R, AG	EQ, C	IA, AG	EQ, C	C	IA	n	n	n	n
10	(250,150)	vs.	(330,70)	IA	IA, R, EQ, AG	IA	IA, EQ, AG	C	C	n	n	n	y

IA = Inequity-aversion, R = Reciprocity, AG = Anti-Greed, EQ = Equity rule, EF = Efficiency rule, ERC = Bolton-Ockenfels (in 3P), C =Competitiveness.
JPM = Is there a joint payoff maximizing allocation available in the respective game? Pareto = Is there a Pareto-dominant allocation available in the respective game? Strict = Is there a strictly equal allocation available in the respective game? Equity = If the less equal allocation is chosen, has the second party a lower payoff than A?

4.4 Experimental Results

We start this section with the analysis of third and second party punishment on an aggregate level. While this provides first insights, the major part of this section is, however, devoted to the analysis of third and second party punishment on the *individual* level, where we present a classification procedure to thoroughly study the driving motivations behind third and second party punishment. We finish this section with an analysis and comparison of the strength of third and second party punishment.

4.4.1. Third and Second Party Punishment: Aggregate Analysis

We observe frequent punishment in both treatments. In 3P, 54 percent of the third parties punish at least once. Furthermore, third parties spend on average 12.7 points per game to punish, more precisely, 8.6 and 4.1 points on the first and the second party, respectively. Table 3 summarizes the frequency and strength of third party punishment in each allocation of each game, distinguishing between punishment for first and second parties.

Table 3— FREQUENCY AND STRENGTH OF PUNISHMENT
THIRD PARTIES

				First Party				Second P. (By-stander)			
Game				Left		Right		Left		Right	
1	(150,150)	vs.	(590,60)	.06	(0.3)	.44	(14.7)	.09	(0.5)	.04	(0.3)
2	(100,100)	vs.	(50,530)	.11	(2.9)	.06	(0.4)	.04	(0.9)	.26	(9.3)
3	(560,60)	vs.	(120,140)	.45	(14.7)	.07	(0.8)	.06	(0.3)	.15	(1.5)
4	(150,90)	vs.	(50,630)	.29	(3.8)	.07	(1.2)	.04	(0.7)	.26	(6.9)
5	(220,260)	vs.	(220,400)	.24	(3.2)	.09	(0.9)	.13	(1.5)	.22	(5.5)
6	(280,240)	vs.	(390,240)	.22	(3.6)	.33	(7.7)	.11	(0.9)	.13	(1.0)
7	(250,80)	vs.	(80,250)	.29	(6.6)	.02	(0.1)	.02	(0.1)	.24	(4.1)
8	(100,100)	vs.	(50,150)	.06	(0.4)	.04	(0.5)	.06	(0.4)	.18	(2.0)
9	(250,150)	vs.	(110,290)	.26	(5.0)	.06	(1.3)	.02	(0.4)	.22	(4.5)
10	(250,150)	vs.	(330,70)	.26	(5.0)	.44	(12.8)	.11	(0.5)	.04	(0.1)

Note: Average points spent for punishment by all participants in parentheses. The endowment of the third party is always 200 points.

In 2P, 60 percent of the second parties punish at least once. Second parties spend on average 13.8 points per game to punish. Table 4 illustrates the frequency and strength of second party punishment in each allocation of each game. We find that the pattern of actual punishment is very well anticipated in 3P and 2P (recall that we elicited the punishment expectations of second and third parties). For instance, there are eight games in which both third and second parties significantly punish the first party more strongly in one allocation ($p<0.05$; Wilcoxon-Signed Rank-Test), and this is anticipated by the first parties in each of these eight games ($p < 0.01$; Wilcoxon-Signed Rank-Test) (see also Figures A and B in the appendix). The behavior of the first parties in 3P and 2P, which is not the focus of our study, can be seen in Table A in the appendix.

Table 4— FREQUENCY AND STRENGTH OF PUNISHMENT
SECOND PARTIES

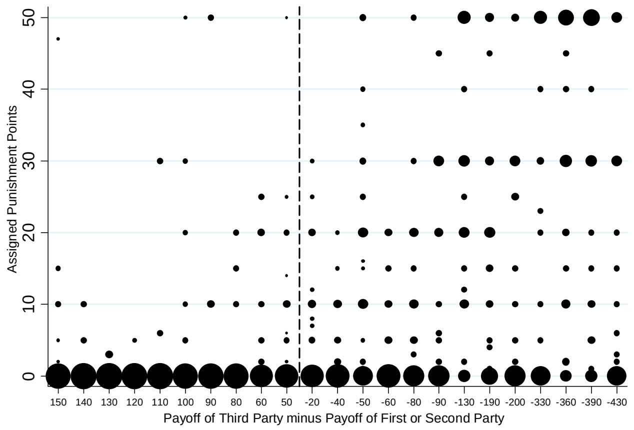
	Game			Left		Right	
1	(150,150)	vs.	(590,60)	.02	(0.2)	.42	(14.7)
2	(100,100)	vs.	(50,530)	.18	(4.1)	.11	(2.3)
3	(560,60)	vs.	(120,140)	.31	(10.3)	.13	(2.9)
4	(150,90)	vs.	(50,630)	.40	(9.6)	.16	(2.7)
5	(220,260)	vs.	(220,400)	.40	(10.6)	.16	(4.0)
6	(280,240)	vs.	(390,240)	.31	(8.2)	.36	(12.7)
7	(250,80)	vs.	(80,250)	.38	(9.1)	.16	(2.9)
8	(100,100)	vs.	(50,150)	.07	(1.7)	.16	(2.6)
9	(250,150)	vs.	(110,290)	.40	(13.6)	.13	(4.0)
10	(250,150)	vs.	(330,70)	.31	(8.1)	.47	(14.3)

Note: Average points spent for punishment by all participants in parentheses.

RESULT 1: *In an aggregate analysis, we observe a strong relationship between disadvantageous inequity and the occurrence of third party punishment. For second parties, we find a strong relationship between disadvantageous inequity as well as harm (reciprocity) and second party punishment.*

Evidence for Result 1: In 3P, the inequity-aversion theories of Fehr and Schmidt (1999) and Falk and Fischbacher (2006) predict that third parties punish another player only if she gets a larger material payoff. This prediction can be reconciled with the data from 14 out of the 15 allocations where a first or a second party is punished by more than 20 percent of the third parties (and/or their average punishment is larger than 3 points). Figure 1 illustrates the individual punishment decisions from the third parties who punish at least once dependent on the size of the payoff differences. More precisely, the horizontal axis indicates the payoff difference between the third and the punished party (the dashed vertical line corresponds to zero distance; to the right of this line the third party gets less than the punished party). The location of the dots indicates the number of points spent from third parties depending on the distance. The size of the dots is proportional to the number of observations, i.e., given a payoff distance between another player and the third party, a dot becomes larger as more and more third parties spend the same amount to punish that player.[57] We can see that payoff differences appear to play an important role in the decision to punish. In fact, third parties rarely punish if their payoff is higher than that of the first/second party, but often and severely if their payoff is lower.

Figure 1: Third Party Punishment depending on Payoff Differences



Explanations other than inequity-aversion seem to play a much less important role in the occurrence of third party punishment. To start, note that reciprocity cannot play any role in 3P because the third party is unaffected by the first and second party, i.e. cannot be harmed. In turn, the model by Bolton and Ockenfels (2000) cannot rationalize much of the observed third party punishment, as it predicts no punishment in games 1 (150/150 vs. 590/60), 3 (560/60 vs. 120/140), 4 (150/90 vs. 50/630) 7 (250/80 vs. 80/250), 9 (250/150 vs. 110/290), and 10 (250/150 vs. 330/70), where we observe considerable punishment. Spite alone (an important ingredient of Levine, 1998) cannot account for the high variation in the frequency of punishment across allocations. It also seems that third parties do not punish deviations from a social norm. For instance, the evidence from game 6 (280/240 vs. 390/240), where we observe frequent punishment of the first player in *both* allocations, suggests that third parties do not punish deviations from norms like an equity norm or a social efficiency norm.

In 2P, we frequently observe considerable punishment in the allocations where it is either predicted by inequity-aversion or reciprocity theories. In many allocations, the behavior points to the importance of inequity-aversion. In game 6 (280/240 vs. 390/240), for instance, about one third of the second parties punish the first party *in either allocation,* even though the first party's choice did not harm them, i.e. the second party's payoff is the same in either allocation.

[57] If there is more than one allocation with a certain payoff difference, we weight the behavior in these allocations uniformly.

There is also considerable punishment in both allocations of game 10 (250/150 vs. 330/70), even when the choice (250/150) is "kind" towards the second party. However, reciprocity also seems important. For instance, inequity-aversion predicts that second parties should never punish in game 5 (220/260 vs. 220/400), while reciprocity predicts punishment in the allocation (220/260). Consistent with this, 40 percent of the second parties punish considerably in the allocation (220/260). There is also some punishment of the allocation (100/100) in game 2 (100/100 vs. 50/530). Other explanations seem not to add much to the understanding of punishment in 2P (however, spite seems important to explain why there is some punishment in most allocations; the individual analysis will clarify this point). In fact, average punishment in any 2P game allocation is stronger than 4 points *if and only if* it is predicted by inequity-aversion or reciprocity.[58] ▪

4.4.2. The Occurrence of Third and Second Party Punishment: Individual Analysis

The previous section provided an aggregate and therefore rather imprecise picture of the motives behind third and second party punishment. We now turn to a more precise analysis on an individual level and provide answers to important questions like: Do third and second parties follow any consistent behavioral patterns? Can we classify the punishers into different types? Which *parsimonious* theory fits our data best? For this, we use the classification procedure from El-Gamal and Grether (1995). This procedure has the crucial advantage that it circumvents the multicollinearity problems that would appear in a classical regression analysis if the decision rules entered as independent variables. Moreover, it allows appropriate inferences even when testing all possible decision rules –no matter how similar their predictions are– at the same time.[59]

The procedure posits that third and second parties follow deterministic decision rules which may differ from subject to subject, but also that they tremble with probability $\varepsilon > 0$, in which case their behavior is random. By selecting the decision rule that best fits each subject's behavior, we can classify subjects in types. Further, we can also find the best single decision rule in 2P and 3P, or the combination of two, three, etc. decision rules that best account for the

[58] The only exception is game 8. In our companion paper we provide a possible explanation, based on the alleviating effect that the strict equality of payoffs has on punishment.

[59] Multicollinearity problems may occur as soon as decision rules share predictions in some allocations (a very common thing in our games). For instance, this is the case for the decision rules that predict no punishment of the second party in 3P and hence share predictions in 20 out of 40 allocations.

behavior in all ten games. Given this, we can then apply the Akaike information criterion to infer the number of decision rules necessary to provide a parsimonious explanation of punishment in our games.

4.4.2.1 Decision Rules in 3P and 2P

In this section, we specify the decision rules that we tested for 3P and 2P. For parsimony, we restrict our analysis to *binary* decision rules, that is, rules indicating only whether the subject punishes and not the strength of punishment. Further, we focus our attention for their particular interest on those binary rules that correspond to the different theories which provide a rationale for costly punishment.[60] Although one could describe the rules for both treatments at the same time, we do it separately for expositional convenience; starting with the formal description of the 2P rules (in section 4.3 we explained the intuition behind most of these rules).

Since second parties in 2P make a total of 20 decisions (one for each of the two allocations in each of the ten games), a decision rule in 2P consists of a vector of 20 ones and zeros: It takes value one if the rule predicts punishment at the corresponding allocation and zero if it predicts no punishment. Thus, there are in principle 2^{20} possible binary decision rules in 2P. For simplicity, however, we focus on nine decision rules in 2P. Letting $(x_{FP}^{L},\ x_{SP}^{L})$ refer to the left-hand and $(x_{FP}^{R},\ x_{SP}^{R})$ to the right-hand allocation at any game (with FP denoting first party and SP denoting second party), they are defined as follows: (1) the "selfish" rule consists of a vector of 20 zeros and predicts never punishment, (2) the "inequity-aversion" rule predicts punishment only at those allocations where $x_{FP}^{i} > x_{SP}^{i}$, for $i = (L,\ R)$, (3) the "reciprocity" rule predicts punishment at any allocation $j = (L,\ R)$ such that $x_{SP}^{j} < x_{SP}^{i}\quad (j \neq i)$, (4) the "spite" rule consists of a vector of 20 ones, (5) the "anti-greed" rule predicts punishment if $x_{FP}^{i} > x_{FP}^{j}$, for $i = (L,\ R)$, (6) the "efficiency" rule predicts punishment in allocation $(x_{FP}^{j},\ x_{SP}^{j})$ only if $x_{FP}^{j} + x_{SP}^{j} < x_{FP}^{i} + x_{SP}^{i}\ \ (j \neq i)$, (7) the "equity" rule predicts punishment in allocation $(x_{FP}^{j},\ x_{SP}^{j})$ only if $\left|x_{FP}^{j} - x_{SP}^{j}\right| < \left|x_{FP}^{i} - x_{SP}^{i}\right|\ \ (j \neq i)$, (8) the "maximin" rule (inspired by Charness and Rabin, 2002) predicts punishment in allocation $(x_{FP}^{j},\ x_{SP}^{j})$ only if

[60] We do not report the complete analysis here. For instance, we tested a large number of "hybrid" decision rules like an "inequity-aversion and reciprocity" rule. Including such "hybrid" rules did not significantly improve the model. The results are available upon request.

$\min\{x_{FP}^j, x_{SP}^j\} < \min\{x_{FP}^i, x_{SP}^i\}$ $(j \neq i)$,[61] and (9) the 'competitiveness' rule predicts punishment at those allocations where $x_{FP}^i \leq x_{SP}^i$. Table 2 in section 4.3 indicates the predictions of some 2P rules (and some 3P ones) in our ten games.

In 3P, third parties make two different punishment decisions in each of the 20 allocations (they can punish the first and/or the second party). Therefore, decision rules in 3P consist of vectors of 40 ones and zeros. We focus on thirteen decision rules, which correspond to existing theoretical approaches. The first eight of them are based on the 2P rules mentioned above: (1) The "selfish" rule consists of a vector of 40 zeros, (2) the "inequity-aversion" rule is a logical extension of the inequity-aversion rule in 2P predicting punishment of the first *and/or* second party when they have a larger payoff than the third party, (3) the "spite" rule is a vector of 40 ones, the (4) "anti-greed", (5) "efficiency", (6) "equity" and (7) "maximin" rules are defined like in 2P (they never predict punishment of the second party in 3P), and (8) the "competitiveness" rule predicts punishment of the first *and/or* second party when they have a smaller payoff than the third party.[62] Further, (9) the "ERC" rule (Bolton and Ockenfels, 2000) predicts punishment of the first *and/or* the second party in allocation $(x_{FP}^j,\ x_{SP}^j, 200)$ if $400 < x_{FP}^j + x_{SP}^j < 600$, i.e., if the third party can use punishment to bring her relative payoff closer to 1/3, the equitable relative payoff in three-player games.[63] Finally, we include some rules for third parties that are based on the ideas of inequity-aversion and reciprocity; this will allow us to investigate if third parties have different motivations than second parties. They are (10) the "indirect reciprocity" rule (inspired by Nowak and Sigmund, 2005 and Seinen and Schram, 2006) predicting punishment of the first party if $x_{SP}^j < x_{SP}^i$ $(j \neq i)$ and no punishment of the second party, (11) an "envy-active" rule predicting punishment of the first party in the same conditions as the inequity-aversion rule, but no punishment of the second party, i.e. people who follow this rule punish richer players only if they are responsible for the outcome,

[61] In other words, this rule predicts punishment for the first party if she does not choose the maximin allocation, maybe because that constitutes a "maximin norm" transgression. Charness and Rabin (2002) report that dictators are often willing to sacrifice part of their material payoff in order to increase the payoff of all recipients, especially of those with low-payoffs.

[62] We do not include a reciprocity rule because it predicts no punishment in the 3P treatment and hence coincides with the selfish rule.

[63] To see this, let $\sigma = x_{FP}^j + x_{SP}^j + 200$ denote the sum of players' payoffs and hence 200/σ denote the third party's relative payoff, and note that (i) this relative payoff is smaller than 1/3 if $400 < x_{FP}^j + x_{SP}^j$, and (ii) a unit of punishment increases the relative payoff if $\frac{200-1}{\sigma-3-1} > \frac{200}{\sigma} \Leftrightarrow \sigma > 800 \Leftrightarrow x_{FP}^j + x_{SP}^j < 600$. Note that we do not include an ERC rule in our analysis of second party punishment because it shares predictions with the inequity aversion rule.

(12) an "egalitarian" rule (Dawes et al., 2007) that predicts punishment (in our games) of those co-players getting a payoff larger than the *average* one,[64] and finally, (13) the "envy-perspective" rule predicting punishment of the first party if $x_{FP}^{i} > x_{SP}^{i}$, and no punishment of the second party (third parties who follow this rule put themselves in the shoes of an inequity-averse second party).

4.4.2.2 Estimation of the Error Rate

The classification procedure posits that each subject follows one of the above mentioned decision rules but allows for mistakes. More precisely, subjects may tremble in each allocation with probability $\varepsilon > 0$, in which case it is assumed that they randomize with equal probability between punishing or not punishing.[65] Consequently, the probability that a subject *s* deviates from her rule at any allocation is $\frac{\varepsilon}{2}$, while the probability of observing her actual behavior - assume that she follows her rule X_s times out of her *d* choices (20 in 2P, 40 in 3P)- is:

$$(1-\frac{\varepsilon}{2})^{X_s} \times (\frac{\varepsilon}{2})^{d-X_s}.$$ [66]

To find the maximum likelihood estimate $\hat{\varepsilon}$ of the error rate, consider first the simplest case: All subjects follow the same decision rule. In that case, $\hat{\varepsilon}$ maximizes the overall likelihood across all *n* players

$$\max \prod_{s=1}^{n} (1-\frac{\varepsilon}{2})^{X_s} \times (\frac{\varepsilon}{2})^{d-X_s} \quad . \qquad (1)$$

One can then prove by applying standard optimization techniques (consult the appendix) that $\hat{\varepsilon}$ coincides with twice the proportion of overall deviations, that is,

[64] More generally, this motive predicts punishment if that reduces the standard deviation from the group mean. This coincides with inequity aversion in two-player games and in many three-player games. In our 3P games, inequity aversion and the egalitarian motive can be discriminated in game 8. While inequity-averse third parties should not damage the bystander in the allocation (50/150), third parties following the egalitarian rule should do so. We find some support for the egalitarian rule since 18 percent damage the bystander then. However, in the classification analysis for third parties the inequity aversion rule outperforms the egalitarian rule.

[65] To simplify the analysis, we assume that all subjects tremble with the same probability in any allocation. This is probably a realistic assumption in view that the punishers' decision problem is, from a strategic point of view, undemanding, so that no change of ε through time (due to learning effects) should be expected.

[66] In computing this, we posit that choices across allocations and games are independent –i.e., the probability of following the rule at any allocation does not depend on what the subjects did before. This seems reasonable in our experiment because (1) subjects are given no feedback and hence there appears to be no reason for changes in mood, and (2) since the punisher's decision problem is arguably easy, we do not expect any learning effects.

$$\hat{\varepsilon} = \frac{2 \cdot (d \times n - \sum_{s} X_s)}{d \times n}. \qquad (2)$$

By computing $\hat{\varepsilon}$ for every possible rule of each treatment, we can then find the optimal decision rule in the maximum likelihood sense, i.e. that maximizing function (1) given the data. This procedure can be extended to the case where different agents use different rules. If we assume that there are two types of players, for instance, we can find the optimal pair of rules by applying the following three-step algorithm to any pair of possible rules A and B: (a) We assign each individual *s* to the rule that minimizes the number of actual deviations $d - X_s$ (in case of a tie, we assign "half" of an individual to each rule), (b) we use expression (2) and the experimental data to find $\hat{\varepsilon}$, and (c) we compute the probability that our data has been generated by the partition of the players generated in step (a), that is,

$$\prod_{i \in A} (1 - \frac{\varepsilon}{2})^{X_i} \times (\frac{\varepsilon}{2})^{d - X_i} \cdot \prod_{j \in B} (1 - \frac{\varepsilon}{2})^{X_j} \times (\frac{\varepsilon}{2})^{d - X_j} \quad . \qquad (3)$$

The optimal pair of rules maximizes equation (3). Finally, if we assume that our subject pool follows three or more rules, the procedure applies analogously.

4.4.2.3 Results of the Classification Procedure

Tables 5 and 6 summarize the results of the classification procedure in 2P and 3P. The second column in each table indicates the best single rule in that treatment, the best pair of rules, and so on. The third column indicates the percentage of second and third parties that follow each rule. The fourth column reports the estimated error $\hat{\varepsilon}$ - recall that the probability that a subject deviates from her rule at any allocation is equal to $\hat{\varepsilon}/2$. Note in this regard that the success of our model (measured by how small $\hat{\varepsilon}$ is) increases as the number of rules *k* increases. This is intuitive as the overall likelihood (3) increases as *k* increases.[67] However, our model also becomes more complex as *k* increases and hence it would be desirable to introduce a penalty for allowing "too many" decision rules. To provide an indication of the optimal number of rules in each treatment, the fifth column of each table reports the log-likelihood – for the best two

[67] The same logic applies here as in a linear regression model, where the coefficient of determination R^2 increases with the number of independent variables.

rules, for instance, this is the log value of (3) – less the number of parameters $(d + n)\cdot k$.[68] According to the Akaike information criterion (AIC), the optimal model should maximize this number. Finally, the sixth column of each table reports the results from a likelihood ratio test of goodness of fit, described later.

TABLE 5— RESULTS OF CLASSIFICATION PROCEDURE IN 3P (THIRD PARTIES)

Number of rules	Rule(s) chosen	Percentage of third parties	ε	AIC	Chi-squared (p-Value)
1	selfish	100%	0.309	-1042.2	
2	selfish, inequity-aversion	78%, 22%	0.182	-862.5	549.31 (0)
3	selfish, inequity-aversion, envy-perspective	66%, 20%, 14%	0.162	-905.6	653.08 (0)
4	selfish, inequity-aversion, envy-perspective, spite	66%, 16%, 14%, 4%	0.151	-971	712.24 (0)

RESULT 2: *A combination of inequity-averse and selfish types can sufficiently capture third parties' punishment patterns. If we allow for more complexity, a combination of two different inequity-averse types, selfish and spiteful types best explains the third parties' punishment pattern.*

Evidence for Result 2: As Table 5 shows, the classification procedure detects the following behavioral patterns for third parties: (1) if we force the algorithm to choose only one rule, the selfish rule is picked. A large number of subjects *never* punish and hence the selfish rule fits their behavior perfectly, and the error rate is already considerably small (0.309 in 3P). The error rates of all other rules are at least twice as high (e.g. inequity-aversion rule: 0.769, envy-perspective rule: 0.667, spite rule: 1). (2) If we force the algorithm to choose the best pair of rules, it selects the selfish rule together with the inequity-aversion rule. Then 22 percent of the third parties are classified as inequity-averse and the error rate drops to 18.2 percent.[69] (3) Adding a third rule is suboptimal according to the AIC, which suggests that the assumption that there are just selfish and inequity-averse types can sufficiently capture the punishment pattern

[68] In a model with k rules, we must first estimate each rule, which consists of *d* zeros and ones (hence the number $d\cdot k$) and moreover we have to find the rule each subject follows or those he or she does not follow (hence the number $n\cdot k$).

[69] In comparison, El-Gamal and Grether (1995) study decisions under uncertainty and find an error rate of 0.312 when looking for the best pair of decision rules.

in 3P. (4) If we nevertheless add a third rule, the algorithm picks the envy-perspective rule. (5) If we add a fourth rule, spite is chosen. ▪

Although the AIC recommends not introducing a third rule if parsimony is our main goal, the comparatively good performance of the envy-perspective rule is an illustration of how this classification procedure can be used to provide new intuitions on punishment.[70] From our knowledge, no experimental paper has provided evidence on this rule before. We speculate that the third parties who follow this rule might be motivated to alleviate the distress of the poorest, weakest party in case that party cannot defend herself (i.e., if she is passive). More experimental evidence, in any case, is required for a better understanding of this kind of behavior.

RESULT 3: *A combination of inequity-averse and selfish types can sufficiently capture second parties' punishment patterns. If we allow for more complexity, a combination of inequity-averse, selfish, spiteful and reciprocal types best explains second parties' punishment patterns.*

Evidence for Result 3: Table 6 indicates the following behavioral patterns for second parties: (1) If we force the algorithm to choose only one rule, the selfish rule is picked. This happens because a large number of subjects *never* punish and hence the selfish rule fits their behavior perfectly. The error rate of 0.502 is therefore quite small compared to that of other rules. The second lowest error rate comes from the inequity-aversion rule which is 0.731, the error rate of the reciprocity rule is 0.798, and the error rate of any other rule is 1. (2) If we force the algorithm to choose the best pair of rules, it selects the selfish rule together with the inequity-aversion rule. We can also see that a considerable fraction of 42 percent is then best classified as inequity-averse. Moreover, we observe that when using these two rules, the error rate is rather low (29 percent). (3) Adding a third rule is suboptimal according to the Akaike information criterion, which suggests that the punishment pattern of second parties can be sufficiently captured by the assumption that there are just selfish and inequity-averse types. (4) However, if we add a third rule, the algorithm picks the spite rule, and 29 percent of the second parties are now classified as inequity-averse and 13 percent as spiteful. (5) If we add a fourth rule, reciprocity is chosen. ▪

[70] Subjects following this rule punish as an inequity-averse second party would do in 2P. For this reason, one might be tempted to think that they just misunderstood the experimental instructions and thought that they were second parties. This is very unlikely, though, as their screens always indicated that they were third parties and they had to indicate their punishment for the first *and* the second party at each allocation.

TABLE 6— RESULTS OF CLASSIFICATION PROCEDURE IN 2P (SECOND PARTIES)

Number of rules	Rule(s) chosen	Percentage of second parties	ε	AIC	Chi-squared (p-Value)
1	selfish	100%	0.502	-572.2	
2	selfish, inequity-aversion	58%, 42%	0.290	-503.4	267.52 (0)
3	selfish, inequity-aversion, spite	58%, 29%, 13%	0.220	-506.9	390.66 (0)
4	selfish, inequity-aversion, spite, reciprocity	53%, 22%, 13%, 11%	0.193	-545.9	442.58 (0)

We note that the results of our classification analysis are in line with Charness and Rabin (2002, p. 838) who suggest that, considering distributional preferences alone (i.e., no reciprocity) and when no self-interest is at stake, approximately 20 percent of their observed behavior can be attributed to difference (i.e., inequity) aversion and 10 percent to spite.[71] Further, and although Dawes et al. (2007) and Falk at al. (2005) do not perform any classification analysis, they cite some behaviors that seem to be explainable only by spite and which have a frequency in line with our previous results. Thus, around 15 percent of the participants in one treatment of Falk at al. (2005) defected *and* punished other players in a Prisoner's Dilemma with a punishment stage (the authors also show that this kind of punishment is very sensitive to its cost; the result that we have just cited corresponds to a punishment technology that is similar to the one available in our experiment).

Observe that the Akaike criterion suggests in both treatments that a model with two, three or four rules is better than one with just one single rule. To further clarify this point, we performed a likelihood ratio test to contrast the null hypothesis that a restricted model with only one rule fits the data similarly well as an unrestricted model with 2, 3, and 4 rules. From the table, we see that we always very strongly reject the null hypothesis.[72]

To sum up, our classification analysis shows that a model assuming two types of players (selfish and inequity-averse) best explains the occurrence of punishment in our two treatments, while alternative and *equally parsimonious* models perform worse. This does not mean, of course, that inequity-aversion can account for the occurrence of all punishment in our games: As we have seen in section 4.4.1, reciprocity plays also an important role in 2P, and other minor variables like spite affect third and second party punishment. Further, the fact that the

[71] Note: Charness and Rabin (2002, p. 823) use the term 'competitive preferences' to refer to what we call 'spite'.

[72] Since negative twice the log-likelihood ratio is asymptotically distributed as chi-squared with degrees of freedom equal to the number of restrictions, large values of the chi-squared statistic reject the null hypothesis. Note that the number of restrictions is d, 2d, and 3d as we restrict 1, 2, and 3 rules, respectively, to coincide with another rule.

error rate ε is never zero indicates that many punishers do not follow strictly a simple decision rule, but take several factors into account when deciding whether to punish.

4.4.3 The Strength of Third and Second Party Punishment

The disadvantage of the classification procedure is that, due to complexity, it makes more sense to investigate the occurrence of punishment only and abstract from its strength. While this is not a problem when testing most theoretical models, we may lose some information concerning models of inequity-aversion and reciprocity which respectively forecast a positive relation between the strength of punishment and the difference in payoffs and the size of the harm, i.e. the net payoff loss of the second party. We first take a look at third parties. A Tobit analysis shows that their average punishment significantly ($p < 0.001$) increases by 7.24 (3.65) points when the difference in payoffs between the first (second) and the third party increases by 100 points (recall that each point spent reduces the payoff of the punished party by 3 points).[73] That is, the bigger the difference in payoffs, the more the third party punishes the first and second parties.

In turn, table 7 shows the results of three Tobit regressions that show whether the difference in payoffs and the size or existence of harm predict the strength of second party punishment. Column (1) reports that, considered in isolation, the difference in payoffs and the size of harm both predict the strength of punishment as suggested, but also that the coefficient for the difference in payoffs is more robust and twice as large as the coefficient for the size of harm. The coefficient of 0.0293 for difference in payoffs means that second parties spend on average 0.0293 points for a payoff disadvantage of one point, i.e. 2.93 points for a payoff disadvantage of 100 points. In column (2), we use the difference in payoffs and the size of harm at the same time in one regression. We can see that when controlling for the difference in payoffs, the size of the harm becomes insignificant. The coefficient for the difference in payoffs remains substantial; the amount of points spent by second parties to punish first parties increases by an average of 2.79 points when the payoff disadvantage increases by 100 points. We also investigate the effect of the sole *existence of harm* by itself. In column (1), we see that the dummy for the existence of harm is a highly significant predictor for the size of punishment when considered in isolation. Second parties are willing to spend 9.58 additional points to punish if they have been harmed. Further, column (3) indicates that the existence of harm alone is also important when we control for the difference in payoffs: The existence of harm then increases punishment by 6.98 points. In summary, subjects punish more if they have been

[73] We use a Tobit regression since players could maximally spend 50 points to punish and many players spent 50 points. An OLS regression analysis leads to very similar results.

harmed, but apparently they do not increase the punishment the more they have been harmed. This leads us to our next result.

TABLE 7—DETERMINANTS OF SECOND PARTY PUNISHMENT (*Tobit)*

Dependent Variable	Strength of Punishment for the First Party		
Column	(1)	(2)	(3)
Difference in payoffs	0.0293*** (0.0062)	0.0279*** (0.0062)	0.0165*** (0.0048)
Size of Harm	0.0138** (0.0058)	0.0039 (0.0058)	
Existence of Harm	9.5813*** (2.6488)		6.9828*** (2.6795)

Notes: 540 Observations. Data comes from all 27 second parties that punish at least once. Data is clustered on individual level. Robust standard errors in parentheses. 58 observations are right-censored at punishment = 50 points. Notes: *** 99-percent significance, ** 98-percent significance, * 95-percent significance.

RESULT 4: *Models that combine envious with reciprocal motives, like Falk and Fischbacher (2006), perform well in predicting the strength of second and third party punishment.*

Evidence for Result 4: The predictions in column (3) are very much in line with Falk and Fischbacher (2006), who predict a relatively more intense punishment of a "richer" first party if she has *also* harmed the second party (independently on the amount of harm inflicted). We observe further support for their theory when comparing games 9 (250/150 vs. 110/290) and 10 (250/150 vs. 330/70). In both games, the first party can choose the allocation (250/150) and this choice leaves the second party in a disadvantageous position (hence some punishment is predicted). In addition, the choice for (250/150) "harms" the second party in game 9, where the alternative allocation is (110/290) but not in game 10 where the alternative is (330/70). As a result, Falk and Fischbacher predict less punishment by second parties of the choice (250/150) in game 10, a prediction which is supported by our data (Wilcoxon-Signed Rank Test, z=2.168, p=0.030).[74] ▪

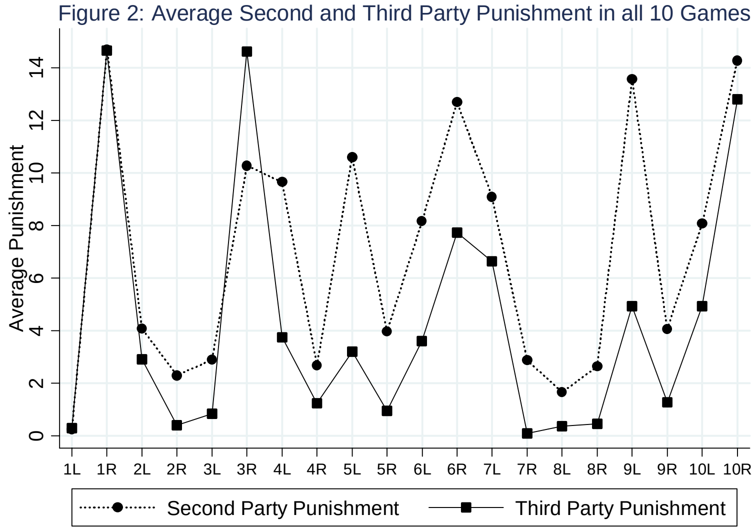
[74] We make two remarks in this regard. First, this characteristic of the model is immaterial in the 3P treatment because third parties are never harmed by any other party. Second, a slightly different version of the model (appendix A of Falk and Fischbacher, 2006) predicts a relatively more intense punishment of a richer first party who harmed the second party *only if* the first party is richer than the second party in the *alternative* allocation. This version thus predicts equal punishment of allocation 250/150 in games 9 and 10, which is not consistent with our data.

4.4.4 Comparing the Strength of Third and Second Party Punishment

RESULT 5*:* *Third party punishment is not generally weaker than second party punishment. Yet, third parties appear to punish more selectively and are especially likely to spend money when their opponents are richer.*

Evidence for Result 5: If we compare how many points third and second parties spend in total, we find no differences (Mann-Whitney Test, z=0.608, p=0.543). This can be explained by two facts: (i) third parties are more sensitive to payoff differences – i.e. they punish a difference of 100 points more than twice as strongly (see section 4.4.3), and, (ii) third parties spend part of their money to punish bystanders. Thus, if we only look at the punishment of the first party, we observe that third parties spend *overall* less points than second parties (Mann-Whitney Test, z=3.209, p=0.001). For a more detailed analysis, figure 2 breaks down what happens in each game. The dots (squares) indicate the average punishment of third (second) parties for the first party in the left- and right-hand allocation in each of the ten games (e.g. 3L = game 3, left-hand allocation) –further, a solid (dotted) line connects the second (third) party observations.[75] Note that the punishment pattern of third and second parties in figure 2 is accurately anticipated by their co-players - figures A and B in the appendix illustrate the average expectation of punishment in each allocation in 3P and 2P. Furthermore, figure 2 illustrates two important findings.

Figure 2: Average Second and Third Party Punishment in all 10 Games



[75] Note that we connected the points in figure 2 just for illustrative reasons. In particular, this does not suggest a temporal ordering. As explained in the experimental design section, the games were played in random order.

First, the solid line lies below the dotted line in most allocations, which indicates that third parties tend to punish the first parties more weakly than the second parties. However, the differences are significant only in 4 of the 20 allocations (Mann-Whitney Test on a 10 percent level: game 5 and 9 left, game 7 and 8 right). For instance, hardly any third party punishes in the allocations (80/250) in game 7 and in (50/150) in game 8 (2 and 4 percent), whereas 16 percent of the second parties punish in these two allocations. Further, only 24 and 26 percent of the third parties punish in the allocations (220/260) in game 5 and (250/150) in game 9 compared to 40 percent of the second parties. These significant differences can be attributed to three reasons: (i) Second parties punish more if they have been harmed, as it happens in these allocations, (ii) second parties tend to be more spiteful than third parties, as our previous classification of third and second parties indicated, and (iii) because payoff differences between first and third parties are rather small in these allocations. *Second,* the figure reveals that third party punishment can be as intense as that from second parties. Remarkably, there are no differences in the allocations that are punished strongest on average (game 1 right: z=-0.054, p=0.956, game 10 right: z=0.327, p=0.743). In the left-hand allocation of game 3, third party punishment is even slightly stronger (z=-1.382, p=0.167). These three allocations have in common that the first party has an income that lies well above the income of the other parties. Hence if large payoff differences exist in our games, second and third parties are equally willing to punish, *even* if the second party has been harmed. ▪

RESULT 6: *Third party punishment very closely resembles second party punishment, which implies that third parties do not act more normatively and impartially than second parties.*

Evidence for Result 6: Figure 2 also illustrates that the pattern of second and third party punishment is identical in all of the ten games. Always, when the second party punished one allocation significantly more strongly than the alternative (which is the case in eight of the ten games), the third party behaved accordingly and punished the same allocation more strongly. This is also the case for third parties in the two remaining games, where second parties punish both allocations approximately equally. This latter fact provides additional evidence that third parties do not punish deviations from a norm of equity, efficiency, or maximin (otherwise they would not punish *both* choices), something supported as well by the classification analysis in section 4.4.2.3.

In addition, the overall similarity in punishment sheds doubt on the assumption that third parties are more impartial and suffer less from an egocentric bias because of their unaffected position in the game (Fehr & Fischbacher, 2004). The behavior in game 7 (250/80 vs. 80/250) also speaks against this assumption. Observe that both allocations in this game are symmetric – i.e., a permutation of each other – so that one should expect that an "impartial" party who uniformly values each player's welfare regards both allocations as equally fair and punish them less than a second party.[76] Contrary to this, we observe that both second and third parties punish the first party equally strongly for choosing the allocation (250/80) (Mann-Whitney Test, z=0.954, p=0.340). ▪

To finish, we address two possible objections to our claim that punishment from third parties is not generally weaker than punishment from second parties. *First*, one might argue that this is an artifact of our setting because second parties have a lower endowment in comparison to third parties in some allocations (regardless of the punishment technology which is the same for third and second parties in all allocations). Indeed, if the marginal utility of money is decreasing in our games, parties with a small endowment should be relatively more reluctant to spend money from their already low endowment. *Second*, the use of the strategy method might have an asymmetric effect on the strength of punishment from third and second parties. In principle, a “hotter” environment induced for instance by the specific response method could increase the strength of punishment from second parties (as in Falk et al., 2005) but not from third parties (since they are unaffected, their reactions might be more independent of the environment).

We can exclude the first objection in our games. Second parties are not more reluctant to spend money if their balance is low. In a Tobit regression analysis, where we use the size of the payoff differences and the endowment of the second party to predict the strength of the punishment in all the cases where the second party endowment is lower than the third party endowment (< 200 points), we find that the second party endowment is an uninformative variable (t=-0.63, p=0.527). In fact, second parties punish especially in games 1, 3 and 10, where their balance is lowest. In summary, the intensity of punishment in our games does not decrease when second parties have a lower balance than third parties.

To address the second objection, we conducted an additional experiment in which both third and second parties played *only* one of our games, now using the specific-response method. We chose game 1 (150/150 vs. 590/60) because we expected a large amount of punishment from our results and also because third and second parties punished the allocation (590/60) equally strongly. The experiment was conducted in Madrid in different university classes, with subjects

[76] We add two remarks. First, in the literature on Welfare Economics, a social welfare function W is said to be symmetric if W(u) = W(u’) whenever the utility vector u constitutes a permutation of vector u’. We have this idea in mind when we refer to impartiality. Second, an impartial spectator might still consider the choice of allocation (250/80) unfair because it fails to be courteous –i.e., that choice signals that the first party cares more for herself than for the second party. Recall, however, that the classification analysis does not find evidence in favor of this anti-greedy behavior.

from different disciplines (60 subjects participated in the 2P and 75 subjects in the 3P treatment).[77] Our data shows that when comparing the behavior of the strategy method with the specific response method, neither second parties nor third parties punish the choice of the allocation (590/60) significantly stronger when using the specific response method (in 3P: z=-1.544, p=0.123); in fact, second parties punish even slightly less when using the specific response method (in 2P: z=1.857, p=0.063). Therefore, this seems to contradict the idea that the specific response method fosters second but not third party punishment in our games.

4.5 Conclusion

We investigate third and second party punishment in a set of ten different games to find out more about the individual motivations behind both types of punishment and to provide insights into the different existing theoretical approaches. The results suggest that inequity-aversion is the crucial (although not the only) cause of third and second party punishment. Our data also shows that third parties do not act more "normatively" or less egocentrically than second parties, casting doubt on the idea that third parties are more impartial.

The evidence from our experiment has implications for the different theoretical models. To start, pure reciprocity models like Rabin (1993), Dufwenberg and Kirchsteiger (2004), and Cox et al. (2007) fail to account for third party punishment. For this reason, we believe that reciprocity alone should not be applied to explain punishment. In contrast, inequity-aversion models like Fehr and Schmidt (1999) and Falk and Fischbacher (2006) fare much better in explaining the occurrence of punishment in 3P and give also rather good predictions in 2P. These models (especially Falk and Fischbacher, 2006) are also more accurate in predicting the strength of punishment. Levine (1998) is inconsistent with the heavy punishment of socially and Pareto efficient actions, and with the role that strict equality plays in reducing punishment (reciprocity also faces this problem). In turn, norm approaches face an unanticipated problem in 2P and 3P: There seems to be no way to explain punishers' choices as a reaction to a prior deviation from any *sensible* norm of distributive justice (taking standard concepts like social efficiency, equity, or maximin into account). A clear illustration of this is that both allocations are punished in some games or that bystanders are damaged by third parties. Indeed, our data indicates that third party punishment is not more "normative". This is not to say, though, that norms are unimportant in explaining punishment, as many third and second parties (even if

[77] The experimental protocol and the instructions were as similar as possible to those of the Zurich sessions. More information on this experiment is available upon request.

inequity-averse) might rationalize their punishment as a reaction to a prior violation of a norm, as the classical philosopher Seneca noted: “Reason wishes the decision that it gives to be just; anger wishes to have the decision which it has given seem the just decision”. People might not punish normatively, but they are likely to believe that they do so.

The results from our classification analysis can be used for predictive purposes. For example, it is a very natural question how a change in the third party endowment could affect punishment. In this regard, our analysis suggests that one group of third parties (the inequity-averse) will probably stop punishing if their endowment rises enough, that is, if they are richer than the other parties. In contrast, other groups (the envy-perspective and spiteful ones) might punish even if they are richer than the other parties (the envy-perspective group would punish only if the first party is richer than the second party). Our evidence provides support in this regard, as we observe less, albeit still some punishment in those allocations where the third party is richer than their co-players (as in some allocations in games 1, 2, 3, 4, and 8), *especially* if in addition the first party is richer than the second party; 29 percent of the third parties punish the first party if she chooses allocation (150, 90) in game 4.

We finish with some possible ideas for future research. First, all the models we have considered deal exclusively with monetary punishment. Hence, none of them is consistent with the idea that people can punish others by non-monetary means (insults, humiliating speech, etc). When do second and third parties use this kind of sanctions and when are they useful in preventing undesirable behavior? Second, since our main objective in this paper was studying and comparing the motives for second and third party punishment, our games have just one sanctioning party. However, it could be interesting to study what happens when there are multiple third parties who can punish, as they might be less willing to punish, on the idea that "others will do it" – this could have to do with the phenomenon of responsibility alleviation reported in Charness (2000). Finally, Falk et al. (2005) report that some defectors in a prisoner’s dilemma punish other players, in particular when the cost of sanctioning is cheap. This correlation between punishment and its cost is not as pronounced for the punishment of defectors by cooperators, and suggests that some type of punishment (spiteful?) is more sensitive to its cost than others (envious, reciprocal?), a topic that deserves also further study.

4.A.1 Appendix Tables and Figures

TABLE A—OBSERVED FREQUENCIES OF CHOICES FROM FIRST PARTIES IN 2P & 3P

	Game			Left		Right	
				2P	3P	2P	3P
1	(150,150)	vs.	(590,60)	.09	.11	.91	.89
2	(100,100)	vs.	(50,530)	.80	.82	.20	.18
3	(560,60)	vs.	(120,140)	1	.98	0	.02
4	(150,90)	vs.	(50,630)	.73	.87	.27	.13
5	(220,260)	vs.	(220,400)	.20	.18	.80	.82
6	(280,240)	vs.	(390,240)	.13	.07	.87	.93
7	(250,80)	vs.	(80,250)	1	1	0	0
8	(100,100)	vs.	(50,150)	.96	1	.04	0
9	(250,150)	vs.	(110,290)	.93	.96	.07	.04
10	(250,150)	vs.	(330,70)	.40	.18	.60	.82

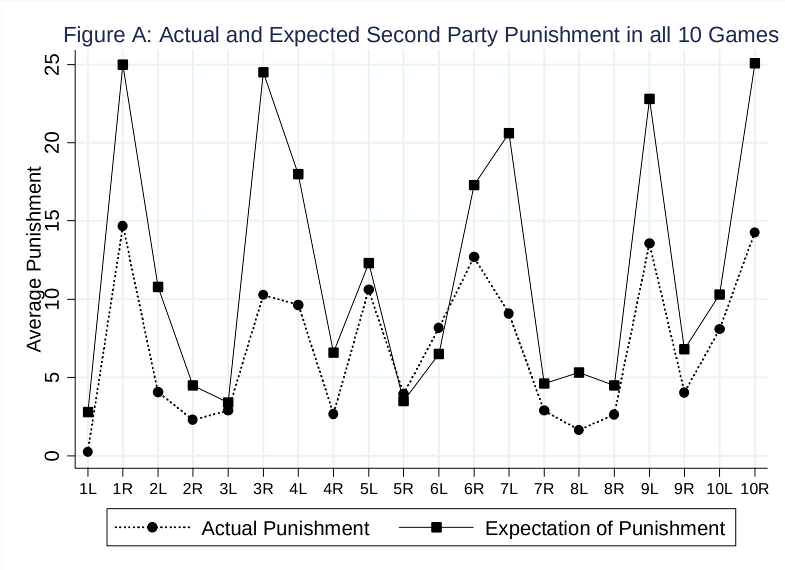


Figure A: Actual and Expected Second Party Punishment in all 10 Games
Average Punishment
0
5
10
15
20
25
1L 1R 2L 2R 3L 3R 4L 4R 5L 5R 6L 6R 7L 7R 8L 8R 9L 9R 10L 10R
Actual Punishment
Expectation of Punishment

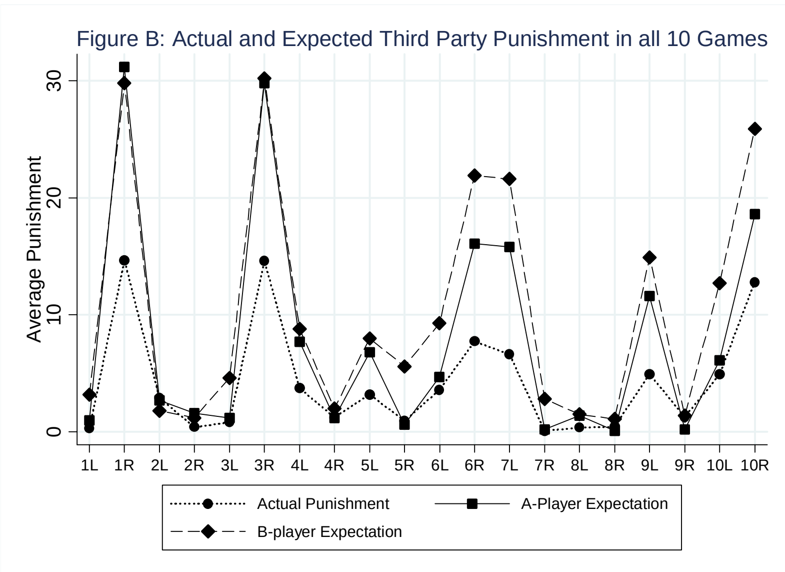


Figure B: Actual and Expected Third Party Punishment in all 10 Games
Average Punishment
0
10
20
30
1L 1R 2L 2R 3L 3R 4L 4R 5L 5R 6L 6R 7L 7R 8L 8R 9L 9R 10L 10R
Actual Punishment
A-Player Expectation
B-player Expectation

4.A.2 Appendix: Derivation of the maximum likelihood estimate of ε.

We obtain $\hat{\varepsilon}$ by solving the following maximization problem (assuming that an interior solution $\hat{\varepsilon} \in (0, 1)$ exists)

$$\max_{\varepsilon \geq 0} \prod_{s=1}^{n} (1 - \frac{\varepsilon}{2})^{X_s} \times (\frac{\varepsilon}{2})^{d - X_s} ,$$

or, equivalently,

$$\max_{\varepsilon \geq 0} \sum_{s=1}^{n} \left[X_s \cdot Ln(1 - \frac{\varepsilon}{2}) + (d - X_s) \cdot Ln(\frac{\varepsilon}{2}) \right]. \qquad (1)$$

Computing the first derivative of (1) and after some algebra, one gets the first order condition of problem (1), that is,

$$\sum_{s=1}^{n} \left[\frac{d(2-\varepsilon) - 2X_s}{\varepsilon(2-\varepsilon)} \right] = \frac{dn}{\varepsilon} - \frac{2\sum X_s}{\varepsilon(2-\varepsilon)} = 0 \quad \Leftrightarrow \quad \hat{\varepsilon} = \frac{2\left[dn - \sum X_s\right]}{dn} \quad , (2)$$

Further, since the second derivative of (1) $-\frac{dn}{\varepsilon^2} + \frac{4(1-\varepsilon)\sum X_s}{[\varepsilon(2-\varepsilon)]^2}$ is negative because $\sum X_s < dn$ and $\frac{4(1-\varepsilon)}{(2-\varepsilon)^2} < 1$, it follows that indeed (2) is a maximum. Finally, and in case expression (2) takes a value equal or larger than 1 so that an interior solution does not exist, the optimum is clearly $\hat{\varepsilon} = 1$. ▪

Appendix: Experimental Instructions

Content of Appendix

In the following I document the English translation for the instructions in Chapter 2, 3 and 4. The experiments in Chapter 2 and 3 were conducted in Brazil and are translated from Portuguese. The instructions for the experiments were presented orally and participants did not receive written instructions. The experiments in Chapter 4 were conducted in Switzerland and Spain and the participants received written instructions. For the experiments in Chapter 4, I only provide the instructions for the Swiss sessions and for one role because the instructions for the Spanish sessions and for the different roles are very similar. The complete set of instructions is available upon request.

A.1 Instructions for Laboratory Experiments in Brazil (Chapter 2 & 3)

A1.1 Instructions for Laboratory Experiments in 2008

Thank you for coming to today's meeting. Please note that you are free to leave this meeting at any point of time. Today's meeting starts with several games. Thereafter, there will be a survey. During the games, you will have the possibility to earn money. The money you earn will be paid out at the end of the meeting. Nobody besides me will know what you will earn today. The payment will be private. You should know that the money comes from research funds and not from our own pockets or from the pocket of politicians. Please note, that there is no right or wrong in playing the games, this is not a test. During today's session you will receive a code. This ensures that everything you do – your decisions in the games and your answers in questionnaires – will remain anonymously.

(Time Preference Experiment, Chapter 2) Before we start, here are two pralines. Please take them. Do you like them? Do you want these two pralines now or do you prefer to wait until the end of today's meeting and receive three instead of two pralines?

We will now play 5 different games. You will be paid according to the outcome from two of the five games. But you will only know after you played all five games for which of the two games you will be paid. After you have played all 5 games, you will draw two cards and the cards will determine which two games will be relevant for payment. This means that you should take your decisions in all five games seriously because there is very high chance that one game will become relevant for your payment. During the five games, we will speak of points. 1 Point is worth 1 Real in the two games that will be chosen for payment. In the other three games, the points will be not converted to Reais.

This is game 1 (Donation Experiment, not reported in Chapter 2 and 3).

I will now give you two envelopes. In the one envelope for you are 10 points, in the other are no points. Your decision is the following: You decide how many of the 10 points you take out of your envelope and transfer to the other. Each point that you transfer from your envelope to the other envelope will be donated to an orphanage. Thus, the more points you take out of your envelope, the less you have, but the more points the orphanage receives. Let me give you two examples: You transfer 9 points; this means you will receive 1 point and the orphanage 9 points. Or, you transfer 3 points which means you will receive 7 points and the orphanage will get 3 points. Of course you can transfer as many of the 10 points as you want, that is, from zero to ten points.

Do you understand? While you make your decision, I will turn my back. Please do not tell me what you plan to do. Please decide now and transfer the amount of points from this envelope to the other envelope and then put the two envelopes in the box in front of you. Tell me when you are ready!

Let's move on to game 2 (Risk-aversion Experiment, Chapter 3).

I will now give you 10 points. They are yours. If this game is one of the two games selected for payment, it would mean that you get 10 Reais. You can play with these points, however, playing is risky: you can multiply these points or lose them. This depends on this coin. You will throw this coin and choose head or tail. If you choose head and head shows up, the points you decided to play with are multiplied by 2.5. If you choose head and tail shows up, you will lose all points you decided to play with. You can decide not to play or to play with 1 – 10 points. Let me give you an example: I decide to play with 5 points, which means that I have 5 points for certain. Then I will choose head or tail and afterwards I will throw the coin. If I choose head and tail shows up, I will only receive 5 points. In contrast, if head shows up I will receive 5 * 2.5 = 12.5 points + 5 points = 17.5 points. Do you understand? How many points do you want to risk?

Now, the third game (Public Goods Experiment Chapter 2 and 3).

The outcome in this game depends on your decisions and the decisions of two others in this meeting. Note that you will never know who these two others are and these two others will never know that they played with you. You and the two others will have to make the same decision. Here are two envelopes. In one envelope, which is denoted *your envelope*, are 10 points. These points are yours. The other, which is denoted *your group envelope*, is empty. You decide how many of the 10 points you transfer to your group envelope. What happens if you transfer points to your group envelope? First, of course, you will have fewer points in your envelope. Second, for every point you transfer to the group envelope, we will add 0.5 points. Thus, if you transfer for example 10 points, we will add 5 points and there will be 15 points in the group envelope. If you transfer nothing, we will not add points to the group envelope. What happens to the points in the group envelope? They will be equally distributed among all participants in your group including you. So, if there are 15 points in the group envelope, you and the other two in your group get 5 points. You do not know how many points the others transfer to the group envelope. The other two participants in your group will also have to decide how many points they transfer to the group envelope before knowing the decisions of their group members.

Let me give you an example. Imagine all three participants (including you) decide to transfer no points to the group envelope. Thus, there are no points in the group envelopes and all three participants stay with their 10 points in their private envelope. Imagine now all three participants including you decide to transfer all 10 points to the group account, i.e., there are 30+0.5x30=45 points in the group envelopes. We will then divide the 45 points equally and each of you will receive 15 points. One last example: Imagine participant 1 gives 10 points to the group envelope, participant 2 gives 0 points to the group envelope und you give 5 points to the group envelope. We will then add 0.5 points for each point in the group envelopes, i.e. there are (10+0+5)x1.5=22.5 points. Then we divide these points equally among the three participants so that all get 7.5 points in addition to the points they kept in their individual envelopes. So, participant 1 gets 0+7,5=7,5 points, participant 2 10+7,5=17,5 points and you 5+7,5=12,5 points. Note that participant 2 received more points than you and participant 1 because he did not transfer any point to the group envelope. In contrast, participant 2 received less because he transferred all 10 points to the group envelope.

Do you understand? While you make your decision, I will turn my back. Please do not tell me what you plan to do. Please decide now and transfer the amount of points you want from this envelope to the other and then put the two envelopes in the box in front of you. Tell me when you are ready!

In this game you have an additional possibility to earn money. You can receive 5 more points if you guess correctly how many points the others transfer to the group envelope. You have two guesses.

This is game 4 (Coordination Experiment, not reported in Chapter 2 and 3).
In this game you will play with one other participant from this meeting but you do not know who, and the other participant does not know that s/he plays with you. You will not know until the end of the meeting how the other participant decided in this game. Imagine you are a hunter. You and the other participant have to make the following decision: Hunting a rabbit or a stag. The rabbit can be hunted individually but hunting the stag is only possible together. If you decide to hunt a rabbit you will get 7 points (= a rabbit is worth 7 points). If both of you decide

to hunt a stag you will both get 10 points (= a stag is worth 20 points). However, if you decide to hunt the stag and the other participant the rabbit, you will get no points (because you cannot hunt the stag alone) and the other will get 7 points (the rabbit can be hunted alone). Likewise, if you decide to hunt the rabbit and the other the stag, you will get 7 points and the other 0 points.

Will you hunt the rabbit or the stag? What do you believe, how will the other participant decide?

This is game 5 (Competition Experiment, not reported in Chapter 2 and 3).
The goal of this game is to throw this tennis ball into the bucket. You have 10 tries. There are two options how you can earn money in this task.

Option 1: For each ball which enters the bucket and stays in it, you will get 1 point.
Option 2:You will play against another fisherman in this meeting but you do not know who. The fisherman who enters more balls in the bucket is the winner. Only the winner gets paid. He gets 3 points for each ball that enters the bucket, the loser gets nothing. In case, you and the other fisherman enter the same amount of balls in the bucket, there is no winner and both of you get 1 point for each ball entered.

Which option do you choose?

A.1.2 Instructions for Laboratory Experiments in 2006

Thank you for coming to today's meeting. As gratification you receive 5 Reais. Today's meeting can take up to 4 hours. Please note that you are free to leave this meeting at any point of time. We ask you to fill out a questionnaire together with one of our helpers. Moreover, we play two games with you. In both games, you have the possibility to earn money. You will be paid at the end of this meeting. Nobody besides me will know how much you will earn today. The payment will be private. You should know that the money comes from research funds and not from our own pockets or from the pocket of politicians. Please note that there is no right or wrong in playing the games. Please do not talk while playing the games. Participants who do not obey this rule cannot continue to play the game. During today's session you will receive a code. This ensures that everything you do – your answers in the questionnaire and your decisions in the games – will remain anonymously.

First Game (Public Goods Experiment, Chapter 2)

In this game you will be in a group of four participants. No participant will know in which group he is. You will have the chance to earn money in this game. How much you will earn depends on your decisions and the decisions of the other three participants in your group. No participant will know how you decided. This experiment has 5 periods. The group composition does not change during these 5 periods. In each period, each group member makes one decision. The decision is the same for all participants. In this experiment, we will speak of points and 3 points are worth 1 Real. At the end of this game, we will add up the points you made during the 5 periods, divide them by three and pay you after you finished playing the second game.

The setting: imagine that you are a fisherman and share a lake with three other fishermen. You all possess a small boat in which there are two fishnets, one fishnet with a small mesh size and one fishnet with a big mesh size.

Your decision: you decide in each of the five periods (which is presented as a week) on how many of the seven days of a week you will use the fishnet with the small mesh size and on how many of the seven days of a week you will use the fishnet with the big mesh size. The fish god "Iara" gives you 10 points in each period. Iara will also give you for every day you will use a fishnet with a big mesh size 0.5 points. For every day you decide to use the fishnet with the small mesh size, Iara will give you 1 point, but at the same time reduce the balance of the other three participants in your group by 0.5 points. You will know after you have made your decision how the other participants decided in the previous period.

Example:

(Explanation for all participants: experimenter is Iara, four helpers stand in front of all participants and represent one group. Experimenter gives each group member ten cards, each representing one point.)

Iara asks the *first* group member: “On how many days of the first week will you use the fishnet with the big mesh size?”
First group member replies: “On *four* days”.
Iara gives the first group member four half cards (each representing one half point)
Iara then gives the first group member three cards (each representing 0.5 points) for the decision to use on the remaining three days with the fishnet with the small mesh size.
Iara approaches all three other group members and cuts three cards from each into two parts and takes half because the first group member decided to catch three days with the fishnet with a small mesh size.

Iara asks the *second* group member: “On how many days of the first week will you use the fishnet with the big mesh size?”
First group member replies: “On *two* days”.
Iara gives the first group member two half cards (each representing 0.5 points)
Iara then gives the second group member five cards (each representing one point) for the decision to use on the remaining three days with the fishnet with the small mesh size.
Iara approaches all three other group members and cuts five cards from each into two parts and takes half because the first group member decided to catch five days with the fishnet with a small mesh size.

Iara asks the *third* group member: “On how many days of the first week will you use the fishnet with the big mesh size?”
First group member replies: “On *seven* days”.
Iara gives the third group member seven half cards (each representing 0.5 points).

Iara asks the *fourth* group member: “On how many days of the first week will you use the fishnet with the big mesh size?”

First group member replies: “On *zero* days”.

Iara gives the fourth group member seven cards (each representing one point) for the decision to use all seven days the fishnet with the small mesh size.

Iara approaches all three other group members and cuts seven cards from each into two parts and takes half because the first group member decided to catch seven days with the fishnet with a small mesh size.

Iara asks the first group member: “Show me your cards!”
First group member replies: “I have 1 card and 16 half cards”
Iara asks the first group member: “So how many points do you have!”
First group member replies: “I have 9 points!”
Iara asks the first group member: “So how much money did you earn in the first week!”
First group member replies: “I made 3 Reais!”

Iara asks the second group member: “Show me your cards!”
Second group member replies: “I have 5 cards and 12 half cards”
Iara asks the second group member: “So how many points do you have!”
Second group member replies: “I have 11 points!”
Iara asks the second group member: “So how much money did you earn in the first week!”
Second group member replies: “I made almost 4 Reais!”

Iara asks the third group member: “Show me your cards!”
Third group member replies: “I have 12 half cards”
Iara asks the third group member: “So how many points do you have!”
Third group member replies: “I have 6 points!”
Iara asks the third group member: “So how much money did you earn in the first week!”
Third group member replies: “I made 2 Reais!”

Iara asks the fourth group member: “Show me your cards!”
Fourth group member replies: “I have 9 cards and 8 half cards”
Iara asks the fourth group member: “So how many points do you have!”
Fourth group member replies: “I have 13 points!”
Iara asks the fourth group member: “So how much money did you earn in the first week!”
Fourth group member replies: “I made a bit more than 4 Reais!”

Questions: Is this clear? We will come to you, explain the game one more time individually, and, ask you about your decision.

Second Game (Time Preference Experiment, Chapter 2 and 3)
The second game is very simple. In this game, you decide whether you want to have a good today or more of the same good tomorrow. If you decide to have the good today, we will give it to you at the end of todays meeting. If you decide to have more of the same good tomorrow, we will give you a voucher to collect the good tomorrow in the village leader's house. Is this clear?

You will decide over three different goods (*the order of the presentation of the three goods was randomly determined*) but you will only receive one of the three goods. After you have made your choice for the three different goods, you will draw a card that determines which of the three goods you will actually receive. Is this clear?

Please tell me now your decision for the following three goods. Do you prefer:
☐ 1,5 L mineral water today or ☐ 3 L mineral water tomorrow
☐ 1,5 Reais today or ☐ 4 Reais tomorrow
☐ 12 bananas today or ☐ 24 bananas tomorrow.

(The decision with regard to money and bananas are not reported in Chapter 2 and 3, the interested reader can consult for further information Fehr and Leibbrandt, 2008).

A.2 Instructions for Laboratory Experiments (Chapter 4)

A.2.1 Instructions for the Second Party in Treatment 2P

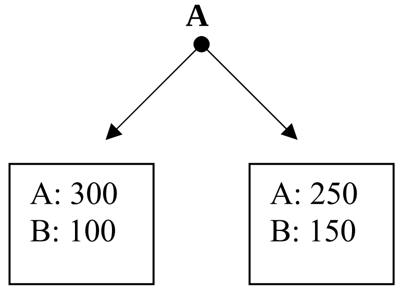
We welcome you to our experiment. If you read the following instructions carefully you will be able to earn money in addition to your show up fee of 10 Swiss Francs – depending on your decisions and the decisions of the other participants. Therefore it is very important, that you read the following instructions carefully. If you have any question, please address them to us.
During the experiment you are not allowed to talk to other participants. If you do not follow this rule we will have to exclude you from the experiment and you will not be able to earn money. In this experiment you will have to make one decision in ten different situations that can influence your payoff. The order of these ten situations is randomly determined. At the end of the experiment, a ten-sided dice will be thrown to determine which of the ten situations becomes relevant for your payment. In this experiment we always speak of points. 10 points are worth 1 Swiss Franc. 10 points = 1 Swiss Franc.

There are two types of participants in this experiment: Participant A and participant B. *You are participant B.* You will never get to know the identity of any participant A (or B), nor will any participant A get to know who you are. The payment at the end of the experiment is also anonymous, that is, no other participant will know how much you earned in this experiment.

Each of the ten situations consists of two stages. In the following, we will explain these two stages.

The first stage

In the first stage, participant A makes his decision. He can decide between two allocations. Take for instance the following example. If he decides for the allocation on the left side, he gets 300 points and you get 100 points. If he decides for the allocation on the right side, he gets 250 points and you get 150 points.



The second stage

In the second stage, you make your decision. You can assign deduction points to participant A. Every deduction point you assign, reduces your payoff by 1 point and the payoff of participant B by 3 points. You can assign between 0 and 50 deduction points. For instance, if you assign 50 deduction points, your payoff is reduced by 50 points and the payoff of participant B by 150 points. If you assign 25 deduction points, your income is reduced by 25 points and the payoff of participant B by 75 points.

Time Line of the Experiment

You will have to decide how many deduction points you assign in all ten different situations before you know which allocation participant A has chosen in the first stage of the situations. We will present you the two different allocations in each situation and you will have to decide how many deduction points you assign in each allocation. While you are making your decisions, participant A will choose one allocation in each situation. After all participants A and B have made their decisions in the ten situations, the experiment is over. One situation will be randomly determined by a ten-sided dice and you will be paid according to your and participant A's decision in this situation.

Calculation of Payoffs

The payoffs of participant A and B are calculated as follows:

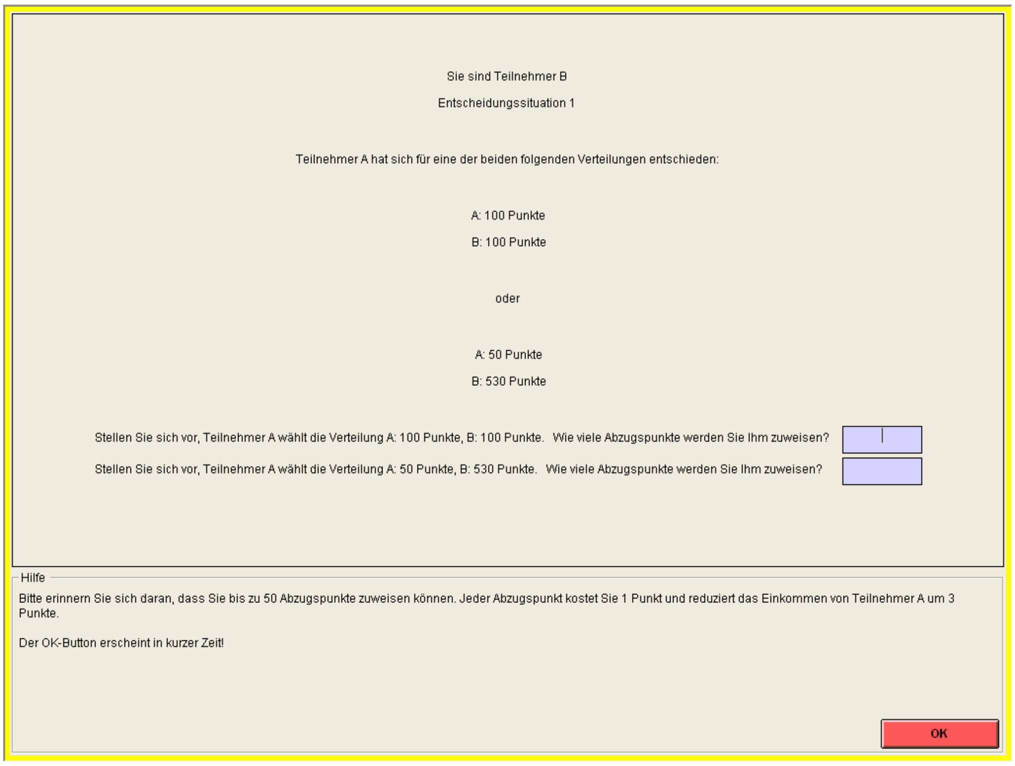
The payoff of participant A =

+ Points for A in the allocation participant A has chosen in the game that was chosen by the dice
– 3x the deduction points you assigned to participant A in the game that was chosen by the dice

Your payoff (participant B) =

+ Points for B in the allocation participant A has chosen in the game that was chosen by the dice
– The deduction points you assigned to participant A in the game that was chosen by the dice

In the second stage, you will have to decide how many deduction points you assign in the first situation. The following computer screen will appear:

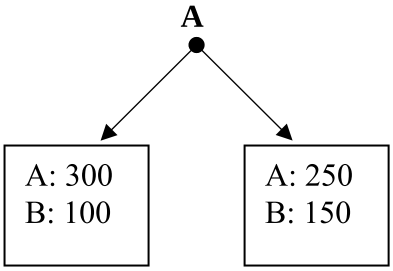


After that, you will make your decision in the second, third,…, tenth situation.
You can take as much time as you need. The OK-Button appears with a little time delay.

If you have made your decisions in all ten situations, you will be informed about your payoff.

Please answer now the following control questions and raise your hand if you have answered them. The experiment starts as soon as all participants have correctly filled out the control questions.

1.

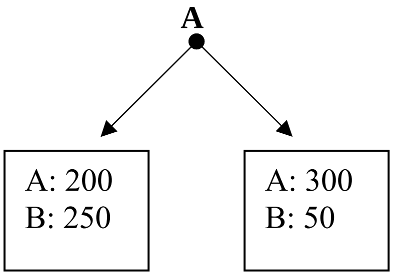


Participant A chooses the allocation on the left side.

a) Participant B assigns 0 deduction points to participant A.
   What is the payoff of participant A? .... B?..........

b) Participant B assigns 30 deduction points to participant A.
   What is the payoff of participant A? .... B?..........

2.



Participant A chooses the allocation on the right side.

a) Participant B assigns 0 deduction points to participant A.
   What is the payoff of participant A? .... B?..........
b) Participant B assigns 50 deduction points to participant A.
   What is the payoff of participant A? .... B?..........

Do you have any further questions?

A.2.2 Instructions for the Third Party in Treatment 3P

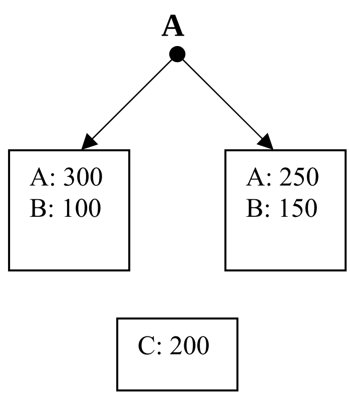
We welcome you to our experiment. If you read the following instructions carefully you will be able to earn money in addition to your show up fee of 10 Swiss Francs – depending on your decisions and the decisions of the other participants. Therefore it is very important, that you read the following instructions carefully. If you have any question, please address them to us.

During the experiment you are not allowed to talk to other participants. If you do not follow this rule we will have to exclude you from the experiment and you will not be able to earn money. In this experiment you will have to make one decision in ten different situations that can influence your payoff. The order of these ten situations is randomly determined. At the end of the experiment, a ten-sided dice will be thrown to determine which of the ten situations becomes relevant for your payment. In this experiment we always speak of points. 10 points are worth 1 Swiss Franc. 10 points = 1 Swiss Franc. There are three types of participants in this experiment: Participant A, participant B and participant C. *You are participant C.* You will never get to know the identity of any participant, nor will any participant get to know who you are. The payment at the end of the experiment is also anonymous, that is, no other participant will know how much you earned in this experiment.

Each of the ten situations consists of two stages. In the following, we will explain these two stages.

The first stage

In the first stage, participant A makes his decision. He can decide between two allocations. Take for instance the following example. If he decides for the allocation on the left side, he gets 300 points and B gets 100 points. If he decides for the allocation on the right side, he gets 250 points and B gets 150 points. Independently of participant A's choice, you will always get 200 points in all ten situations.



The second stage

In the second stage, you make your decision. You can assign deduction points to participant A and/or participant B. Every deduction point you assign to participant A (or participant B), reduces your payoff by 1 point and the payoff of participant A (or B) by 3 points. You can assign in total between 0 and 50 deduction points. For instance, if you assign 50 deduction points to participant A, your payoff is reduced by 50 points and the payoff of participant A by 150 points. If you assign 30 deduction points to participant B, your income is reduced by 30 points and the payoff of participant B by 90 points.

Time Line of the Experiment

You will have to decide how many deduction points you assign in all ten different situations before you know which allocation participant A has chosen in the first stage of the situations. We will present you the two different allocations in each situation and you will have to decide how many deduction points you assign in each allocation. While you are making your decisions, participant A will choose one allocation in each situation and participant B will be asked how many deduction points you will assign. After all participants have made their decisions in the ten situations, the experiment is over. One situation will be randomly determined by a ten-sided dice and you will be paid according to your decision in this situation.

Calculation of Payoffs

The payoffs of participant A, B and C are calculated as follows:

The payoff of participant A =

+ Points for A in the allocation participant A has chosen in the game that was chosen by the dice
– 3x the deduction points you assigned to participant A in the game that was chosen by the dice

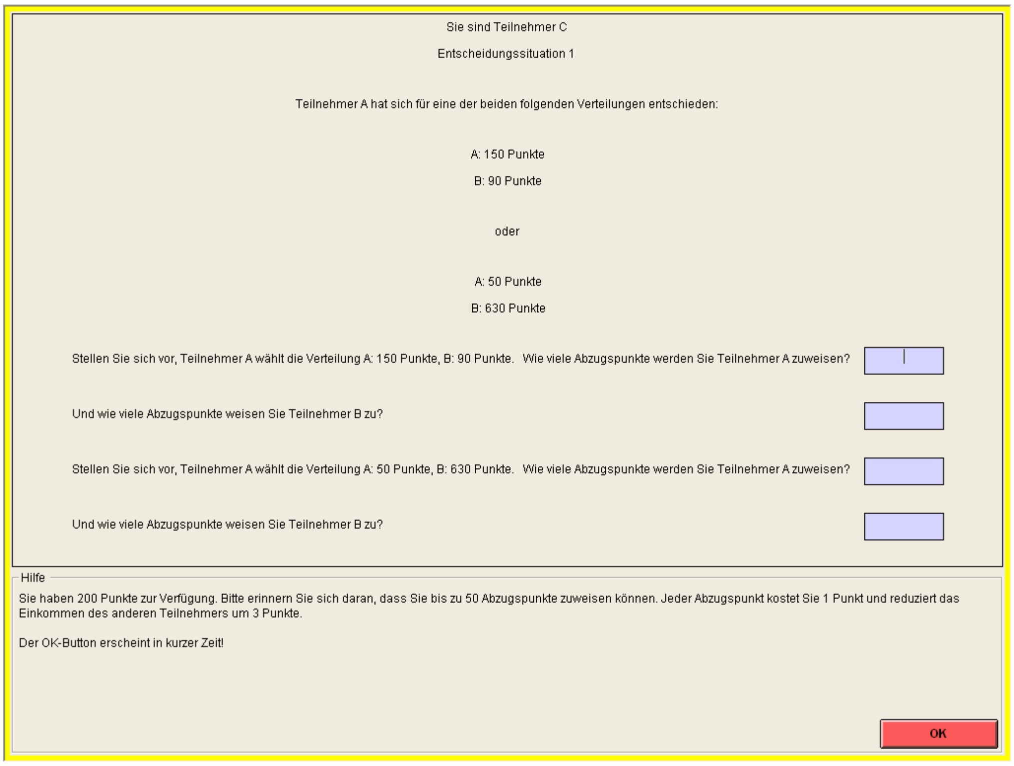
The payoff of participant B =

+ Points for B in the allocation participant A has chosen in the game that was chosen by the dice
– 3x the deduction points you assigned to participant B in the game that was chosen by the dice

Your payoff (participant C) =

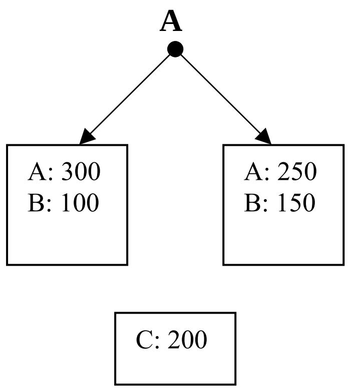
- \+ 200 Points (your endowment)
- – The deduction points you assigned to participants A and/or B in the game that was chosen by the dice

In the second stage, you will have to decide how many deduction points you assign in the first situation. The following computer screen will appear:



After that, you will make your decision in the second, third,…, tenth situation. You can take as much time as you need. The OK-Button appears with a little time delay. If you have made your decisions in all ten situations, you will be informed about your payoff. Please answer now the following control questions and raise your hand if you have answered them. The experiment starts as soon as all participants have correctly filled out the control questions.

1.

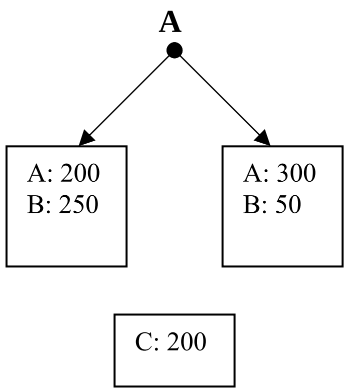


Participant A chooses the allocation on the left side.

a) Participant C assigns 0 deduction points to participant B.
What is the payoff of participant A? .... B?.......... C?............

b) Participant C assigns 30 deduction points to participant B.
What is the payoff of participant A? .... B?.......... C?............

2.



Participant A chooses the allocation on the right side.

a) Participant C assigns 0 deduction points to participant A.
What is the payoff of participant A? .... B?.......... C?............

b) Participant C assigns 50 deduction points to participant A.
What is the payoff of participant A? .... B?.......... C?............

Do you have any further questions?

Bibliography

Andreoni, James. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving?", *Economic Journal*, 1990, 100(401), pp. 464-477.

Andreoni, James. "Cooperation in Public Goods Experiments: Kindness or Confusion?", *American Economic Review*, 1995, 4, pp.891-904.

Ashraf, Nava; Karlan, Dean and Wesley, Yin. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 2006, 121(2), pp. 635-672.

Babcock, Linda, Loewenstein, George, Issacharoff, Samuel, and Camerer, Colin. "Biased Judgments of Fairness in Bargaining", *American Economic Review*, 1995, 85(5), pp. 1337-1343.

Bandiera, Oriana; Barankay, Iwan and Rasul, Imran. "Social Preferences and the Response to Incentives: Evidence from Personnel Data", *Quarterly Journal of Economics*, 2005, pp. 917-961.

Bendor, Jonathan and Swistak, Piotr. "The Evolution of Norms", *American Journal of Sociology*, 106, pp. 1493-1545.

Benz, Matthias and Meier, Stephan. "Do People Behave in Experiments as in Real Life? Evidence from Donations", *Experimental Economics, forthcoming*.

Benzion, Uri; Rapoport, Amnon and Yagil, Joseph. "Discount Rates Inferred From Decisions: An Experimental Study", *Management Science*, 1989, 35, pp. 270-84.

Bolton, Gary E. and Ockenfels, Axel. "ERC: A Theory of Equity, Reciprocity, and Competition", *American Economic Review*, 2000, 90(1), pp. 166-193.

Bonin, Holger; Dohmen, Thomas; Falk, Armin; Huffman, David and Sunde, Uwe. "Cross-sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes", *Labour Economics*, 2007, 14(6), pp. 926-937.

Brandts, Jordi and Charness, Gary. "Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games", *Experimental Economics*, 2000, 2(3), pp. 227-238.

Brown, Martin; Falk, Armin and Fehr, Ernst. "Relational Contracts and the Nature of Market Interaction", *Econometrica*, 2005, 72(3), pp. 747-780.

Camerer, Colin. "Behavioral Game Theory", 2003, Princeton University Press: Princeton.

Camerer, Colin and Hogarth, Robin M. "The Effect of Financial Incentives in Experiments", *Journal of Risk and Uncertainty*, 1999, 19, pp. 7-42.

Camerer, Colin and Thaler, Richard. "Ultimatums, Dictators, and Manners", *Journal of Economic Perspectives*, 1995, IX, pp. 209-219.

Cameron, Lisa. "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia", *Economic Inquiry*, 27, pp. 47-59.

Cardenas, Juan-Camilo. "How Do Groups Solve Local Commons Dilemmas? Lessons from Experimental Economics in the Field". *Environment, Development and Sustainability*, 2, pp. 305-322.

Carpenter, Jeffrey P. and Matthews, Peter H. "Norm Enforcement: Anger, Indignation or Reciprocity?", *IZA Discussion Paper Series* No. 1583, 2005.

Carpenter, Jeffrey and Seki, Erika. "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay", 2005, *IZA Working Paper* No.1697.

Casari, Marco and Plott, Charles. "Decentralized Management of Common Property Resources: Experiments with a Centuries-Old Institution", 2003, *Journal of Economic Behavior and Organization*, 51(2), pp. 217-247.

Cason, Timothy and Mui, Vai-Lam. "Social Influence in the Sequential Dictator Game", *Journal of Mathematical Psychology*, 1998, 42, pp. 248-465.

Cavalcanti, Carina. "Recursos Comuns em Unidades de Conservacao de Uso Sustentavel: Estudo de Caso de uma APA", 2003, *Unpublished Working Paper.*

Cavalcanti, Carina; Schläpfer, Felix and Schmid, Bernard. "Participation and Willingness to Cooperate in CPR Management: A Field Experiment with Fishing Communities in Brazil", 2008, *Unpublished Working Paper.*

Charness, Gary. "Responsibility and Effort in an Experimental Labor Market", *Journal of Economic Behavior and Organization*, 2000, 42(3), pp. 375-384.

Charness, Gary, Cobo-Reyes, Ramón and Jiménez, Natalia. "An Investment Game with Third-party Intervention", *Journal of Economic Behavior and Organization*, 2008, 68(1), pp. 18-28.

Charness, Gary and Grosskopf, Brit. "Relative Payoffs and Happiness: An Experimental Study", *Journal of Economic Behavior and Organization*, 2001, 45, pp. 301-328.

Charness, Gary, and Rabin, Matthew. "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics*, 2002, 117, pp. 817-869.

Charness, Gary, and Villeval, Marie Claire. "Cooperation and Competition in Intergenerational Experiments in the Field and the Laboratory", *American Economic Review, forthcoming.*

Cox, James C., Friedman, Daniel, and Gjerstad, Steven. A Tractable Model of Reciprocity and Fairness, *Games and Economic Behavior*, 2007, 59, 17-45.

Croson, Rachel. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games", 2008, *Working Paper*.

Dawes, Christopher, Fowler, James H., Johnson, Tim, McElreath, Richard and Smirnov, Oleg. "Egalitarian Motives in Humans", *Nature*, 446, pp. 794-796.

DellaVigna, Stefano. "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, forthcoming.

Dohmen, Thomas; Falk, Armin; Huffman, David and Sunde, Uwe. "Homo Reciprocans: Survey Evidence on Behavioral Outcomes", *Economic Journal*, forthcoming.

Dufwenberg, Martin, and Kirchsteiger, Georg. "A Theory of Sequential Reciprocity", *Games and Economic Behavior*, 2004, 47(2), pp. 268-298.

El-Gamal, Mahmoud, and David Grether. "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, 90(432), December 1995, pp. 1137-1145.

Elster, Jon. "Social Norms and Economic Theory", *Journal of Economic Perspectives*, 1989, 3(4), pp. 99-117.

Engelmann, Dirk, and Strobel, Martin. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments", 2004, *American Economic Review*, 94(4), pp. 857-869.

Engle-Warnick, Jim. "Inferring Strategies from Observed Actions: A Nonparametric, Binary Tree Classification Approach", *Journal of Economic Dynamics and Control*, 2003, 27(11-12), pp. 2151-2170.

Falk, Armin. "Gift-Exchange in the Field", *Econometrica*, 75(5), pp. 1501-11.

Falk, Armin; Fehr, Ernst and Fischbacher, Urs. "Driving Forces behind Inform Sanctions", *Econometrica*, 2005, 7(6), pp. 2017-30.

Falk, Armin, and Fischbacher, Urs. "A Theory of Reciprocity", *Games and Economic Behavior*, 2006, 54(2), pp. 293-315.

Falk, Armin and Kosfeld, Michael. "The Hidden Cost of Control", *American Economic Review*, 2006, 96, pp. 1611-1630

Farzin, Hossein. "The Effect of the Discount Rate on Depletion of Exhaustible Resources", *Journal of Political Economy*, 1984, 92(5), pp. 841-851.

Feeny, David; Berkes, Fikret; McCay, Bonnie and Acheson, James. "The tragedy of the commons. Twenty-Two Years Later", *Human Ecology*, 1990, 18(1), pp. 1-19.

Fehr, Ernst and Fischbacher, Urs. “Third Party Punishment and Social Norms”, *Evolution and Human Behavior*, 2004, 25, pp. 63-87.

Fehr, Ernst and Gächter, Simon. “Cooperation and Punishment in Public Goods Experiments”, *American Economic Review*, 2000, 90, pp. 980-94.

Fehr, Ernst and Götte, Lorenz. “Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment”, *American Economic Review*, 2008, 97(1), pp. 298-317.

Fehr, Ernst; Kirchsteiger, Georg and Riedl, Arno. „Does Fairness prevent Market Clearing?“, *Quarterly Journal of Economics*, 1993, 108, pp. 437-460.

Fehr, Ernst; Klein, Alexander and Schmidt, Klaus. “Fairness and Contract Design“, *Econometrica*, 2007, 5 (1), pp. 121-154.

Fehr, Ernst and Leibbrandt, Andreas. “Cooperativeness and Impatience in the Tragedy of the Commons”, *IEW Working Paper* # 378, 2008.

Fehr, Ernst, Näf, Michael, and Schmidt, Klaus. “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment”, *American Economic Review*, 2006, 96, pp. 1912-1917.

Fehr, Ernst and Schmidt, Klaus. “A Theory of Fairness, Competition and Cooperation”, *Quarterly Journal of Economics*, 1999, 114(3), pp. 817-68.

Fehr, Ernst; Tougareva, Elena and Fischbacher, Urs. “Do High Stakes and Competition Undermine Fairness? Evidence from Russia”, *IEW Working Paper # 132*, 2002.

Fischbacher, Urs. “z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, *Experimental Economics*, 2007, 10 (2), pp. 171-178.

Fischbacher, Urs; Gächter, Simon and Fehr, Ernst. “Are People Conditionally Cooperative? Evidence from a Public Goods Experiment”, *Economics Letters*, 2001, 71, pp. 397-404.

Frederick, Shane; Loewenstein, George and O'Donoghue, Ted. “Time Discounting and Time Preference: A Critical Review”, *Journal of Economic Literature*, 2002, 40, pp. 351-401.

Frey, Bruno and Meier, Stephan. “Social Comparison and Pro-Social Behavior: Testing Conditional Cooperation in a Field Experiment”, *American Economic Review*, 2004, 94(5), pp. 1717-22.

Gneezy, Uri and List, John. “Putting Behavioral Economics to Work: Field Evidence on Gift Exchange”, *Econometrica*, 74 (5), pp. 1365-1384.

Greif, Avner. “Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition”, *American Economic Review*, 1993, 83, pp. 525-548.

Greif, Avner. “Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies”, *Journal of Political Economy*, 1994, 102, pp. 912-950.

Greiner, Ben. “An Online Recruitment System for Economic Experiments”, in: *Forschung und wissenschaftliches Rechnen 2003*, ed. by Kurt Kremer and Volker Macho, GWDG Bericht 63, Göttingen, Ges. für Wiss. Datenverarbeitung.

Güth, Werner, Schmittberger, Rolf, and Schwarze, Bernd. “An Experimental Analysis of Ultimatum Bargaining”, *Journal of Economic Behavior and Organization*, 1982, 3, pp.367-388.

Güth, Werner, Huck, Steffen and Müller, Wieland. “The Relevance of Equal Splits in Ultimatum Games”, *Games and Economic Behavior*, 2001, 37, pp. 161-169.

Hardin, Garrett. “The Tragedy of the Commons”. *Science,* 1968, 162 .pp.1246-48.

Homans, George. “Social Behavior”, 1961, New York: Harcourt, Brace & World.

Hoeffler, Steve and Ariely, Dan. “Constructed Stable Preferences: A Look Into Dimensions of Experience and Their Impact on Preference Stability”, *Journal of Consumer Psychology*, 1999, 10(2), pp. 113-139.

Kahneman, Daniel. “Comment on Plott's Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis”, in Arrow, Colombatto, Perleman, Schmidt, eds., *Rational Foundations of Economic Behavior*, 1996, London: Macmillan and St. Martins, pp. 251-254.

Kandori, Michihiro. “Social Norms and Community Enforcement”, *American Economic Review,* 1992, 59(1), pp. 63-80.

Karlan, Dean. “Using Experimental Economics to Measure Social Capital and Predict Financial Decisions”, *American Economic Review*, 2005, 95(5), pp. 1688-99.

Kennedy, Randall. “Race, Crime and the Law”, 1997, pp. 168-255, New York: Pantheon.

Keser, Claudia and van Winden, Frans. “Conditional Cooperation and Voluntary Contributions to Public Goods”, *Scandinavian Journal of Economics*, 2002, 102(1), pp. 23-39.

Kirchsteiger, Georg. “The Role of Envy in Ultimatum Games”, *Journal of Economic Behavior and Organization*, 1994, 25(3), pp. 373-389.

Kritikos, Alexander and Bolle, Friedel. “Distributional Concerns: Equity- or Efficiency Oriented?” *Economics Letters*, 2001, 73, pp. 333-338.

Kutner, Michael; Nachtsheim, Chris and Neter, John. “Applied Linear Regression Models”, 2004, McGraw-Hill.

Laibson, David; Repetto, Andrea and Tobacman. "Estimating Discount Functions with Consumption Choices over the Lifecycle", *American Economic Review*, forthcoming.

Landry, Craig; Lange, Andreas; List, John; Price, Michael and Rupp, Nicholas. "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment", *Quarterly Journal of Economics*, 2006, 121(2), pp. 747-782.

Laury, Susan and Taylor, Laura. "Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good?", *Journal of Economic Behavior and Organization*, 2008, 65, pp. 9-29.

Ledyard, John. "Public Goods: A Survey of Experimental Research" in John H. Kagel and Alvin E. Roth, eds., *Handbook of experimental economics*. Princeton: Princeton University Press, 1995, pp. 111-94.

Levine, David K. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, 1998, 1, pp. 593-622.

Levitt, Steven and List, John. "What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?", *Journal of Economic Perspectives*, 2007, 21(2), pp. 153-74.

List, John. "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions", *Journal of Political Economy*, 2006, 114(51), pp. 1-37.

López-Pérez, Raúl. "Aversion to Norm-Breaking: A Model", *Games and Economic Behavior*, 2008, 64, pp. 237-267.

Mas, Alexandre and Moretti, Encrio. "Peers at Work", *American Economic Review*, forthcoming.

Masclet, David; Noussair, Charles; Tucker, Steven and Villeval, Marie-Claire. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism", *American Economic Review*, 93(1), pp. 366-380.

Meier, Stephan and Sprenger, Charles. "Impatience and Credit Behavior: Evidence from a Field Experiment", 2007, *Federal Reserve Bank of Boston Working Papers 07-3*.

Mischel, Walter; Shoda, Yuichi and Rodriguez, Monica L. "Delay of Gratification in Children", *Science*, 1989, 244, pp. 933-38.

Orne, Martin T. "On the Social Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications", *American Psychologist*, 1962, 17(10),pp. 776-83.

Ostrom, Elinor; Burger, Joanna; Field, Christopher; Norgaard, Richard and Policansky, David. "Revisting the Commons: Local Lessons, Global Challenges", *Science*, 1999, 284, pp. 278-82.

Ostrom, Elinor and Nagendra, Harini. "Insights on linking forests, trees, and people from the air, on the ground, and in the laboratory", 2006, *Proceedings of the National Academy of Sciences*, 103(51), pp. 19221-19223.

Ostrom, Elinor; Walker, James and Gardner, Roy. "Covenants with and without a Sword: Self-Governance is Possible", *American Political Science Review*, 1992, 86(2), pp. 404-417.

Plott, Charles. "Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis", in Arrow, Colombatto, Perleman, Schmidt, eds., *Rational Foundations of Economic Behavior*, 1996, London: Macmillan and St. Martins, pp. 225-250.

Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics", *American Economic Review*, December 1993, 83(5), pp. 1281-1302.

Ross, Lee, Greene, David and House, Pamela. "The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes", *Journal of Experimental Social Psychology*, 1977, 13(3), pp. 279-301.

Roth, Alvin. "*Bargaining Experiments*", 1995, in J.Kagel and A.Roth, eds. Handbook of Experimental Economics, Princeton University Press: Princeton.

Seinen, Ingrid and Schram, Arthur. "Social Status and Group Norms: Indirect Reciprocity in a Repeated Helping Experiment", *European Economic Review*, 2006, 50, pp. 581-602.

Shang, Jen and Croson, Rachel. "Field Experiments in Charitable Contributions: The Impact of Social Influence on the Voluntary Provision of Public Goods", 2008, *Working paper*.

Sneath, David. "State policy and pasture degradation in Inner Asia", *Science*, 1998, 281, pp.1147-48.

Segal, Uzi and Segal, Joel. "Tit for tat: Foundations of preferences for reciprocity in strategic settings", *Journal of Economic Theory*, 2007, 136, pp. 197-216.

Sobel, Joel. "Interdependent Preferences and Reciprocity", *Journal of Economic Literature*, 2005, XLIII, pp. 392-436.

Sobel, Joel. "Markets Make People Selfish", *Working Paper*, 2007.

Slovic, Paul. "The Construction of Preferences", *American Psychologist*, 1995, 50, pp. 364-371.

Thaler, Richard and Benartzi, Shlomo. “Save More Tomorrow (TM): Using Behavioral Economics to Increase Employee Saving”, *Journal of Political Economy*, 2004, 112(1), pp. 164-187.

Thompson, Leigh. “The Mind and Heart of the Negotiator”, 2005, Englewood Cliffs, NJ: Prentice Hall.

Walker, James and Gardner, Roy, and Ostrom, Elinor. “Rent Dissipation in a Limited Access Common-Pool Resource: Experimental Evidence”, *Journal of Environmental Economics and Management*, 1990, 19, pp. 203-211.

Zizzo, Daniel J. “Money Burning and Rank Egalitarianism with Random Dictators”, *Economics Letters*, 2003, 81, pp. 263-66.

Curriculum Vitae

Andreas Leibbrandt was born 07/15/1977 in Rastatt (Germany). He visited the Friedrich-List Gymnasium in Asperg where he passed his high school exam in 1997. Thereafter, he made an apprenticeship in Stuttgart at Raiffeisen- und Volksbanken Insurance Company before he studied Economics, Management, Psychology and Sociology at the University of Erlangen-Nuremberg and McGill University in Montréal. In 2005, he graduated at the University of Erlangen-Nuremberg and achieved the title Magister Artium. Since 2005 he is a doctoral student and research assistant at the Institute for Empirical Research in Economics at the University of Zurich.